



UNIVERSIDAD COMPLUTENSE DE MADRID
Facultad de Informática

Proyecto Fin de Máster en Ingeniería Informática para la Industria

Título: Análisis de técnicas de fusión de imágenes para su uso en fotografía digital

Autor: Daniel Alonso Fernández

Director del proyecto: Gonzalo Pajares Martinsanz

Curso académico: 2009-2010

Máster en Investigación en Informática

Facultad de Informática, Universidad Complutense de Madrid

Resumen

Desde su aparición, el mercado de las cámaras digitales no ha dejado de crecer. Ante el aumento de la oferta, los consumidores se inclinan cada vez más por los productos con funciones adicionales que faciliten su uso. El propósito de este trabajo es estudiar los métodos más relevantes existentes actualmente en relación a la fusión de imágenes para decidir cuál se podría incorporar en una cámara para obtener fotografías con una mayor profundidad de campo y con una calidad de la imagen aceptable.

Palabras clave

Fusión de imágenes, profundidad de campo, cámaras digitales, fotografía.

Abstract

From his appearance, the market of the digital cameras has not stopped growing. With the increase of the offer, the consumers incline increasingly for the products with additional functions that facilitate its use. The intention of this work is to study the most relevant existing methods nowadays related to the image fusion. This is intended in order to decide which one could be embedded in the camera devices for obtaining images with a depth of field and simultaneously with an acceptable image quality.

Keywords

Image fusion, depth of field, digital cameras, photography.

Índice

Resumen.....	3
Palabras clave.....	3
<i>Abstract</i>	3
<i>Keywords</i>	3
Capítulo 1: Consideraciones preliminares	7
1.1. Introducción	7
1.2. Motivación y objetivos	7
1.3. Metodología	8
1.4. Organización del trabajo	9
Capítulo 2: Revisión de métodos de fusión	11
2.1. Introducción	11
2.2. Media aritmética	12
2.3. Media ponderada	12
2.4. Pirámide de diferencia de paso bajo.....	14
2.5. Pirámide Laplaciana FSD	17
2.6. Pirámide de razón de paso bajo	19
2.7. Pirámide del gradiente	20
2.8. Pirámide morfológica	22
2.9. Transformada de wavelets.....	24
2.10. Transformada de wavelets invariante a traslaciones.....	26
2.11. Otros métodos de combinación de coeficientes.....	27
Capítulo 3: Fusión en imágenes de color.....	31
3.1. Espacios de color	31
3.2. Espacios de color para fusión de imágenes	32
Capítulo 4: Fusión en imágenes con dimensiones inadecuadas.....	35
4.1. Usar las dimensiones originales	35
4.2. Añadir bandas negras.....	35
4.3. Redimensionar las imágenes.....	36
Capítulo 5: Resultados	37
5.1. Resultados de los métodos de fusión	37
5.2. Otros métodos de combinación de coeficientes	40
5.3. Fusión en imágenes de color.....	44
5.4. Fusión en imágenes con dimensiones inadecuadas	48

Capítulo 6: Conclusiones y trabajo futuro	53
Anexo I: Altas y bajas frecuencias en imágenes	57
Anexo II: Tablas de resultados.....	61
Bibliografía	73
Autorización	77

Capítulo 1

Consideraciones preliminares

1.1. Introducción

En fotografía definimos la profundidad de campo como la distancia por delante y por detrás del punto enfocado que aparece con nitidez, de forma que los objetos que se encuentren dentro de la profundidad de campo aparecen en la imagen enfocados tanto más cuanto más cerca estén del punto ideal. El resto de objetos situados fuera de dicha profundidad aparecen con un grado de desenfoque gradual que se incrementa a medida que los mismos se alejan de la zona definida por la profundidad de campo.

La profundidad de campo depende de la apertura de diafragma, de la distancia focal y de la distancia real entre la cámara y el punto enfocado (TheWebfoto, 2008).

En algunos dispositivos que incorporan cámaras, tales como los teléfonos móviles, al igual que en las cámaras ultra compactas, las fotografías suelen quedar nítidas tanto en el primer plano como en el fondo. Esto se debe a que en este tipo de cámaras, tanto la apertura como la distancia focal o el punto de enfoque son fijos, por lo que la profundidad de campo es siempre la misma. Además, al ser la longitud focal muy pequeña, en torno a 5 mm, esta profundidad de campo resulta ser relativamente grande.

Sin embargo, en las cámaras de gama superior la longitud focal suele ser mayor, en torno a los 50 mm, lo que hace que la profundidad de campo sea mucho menor. A pesar de que la posibilidad de desenfocar el fondo es un efecto muy deseado en algunos casos, como en la fotografía de retratos, en otras circunstancias el usuario puede sentirse frustrado al no poder obtener fotografías en las que esté enfocado “todo”, esto es, el conjunto de objetos presentes en la escena.

1.2. Motivación y objetivos

En este trabajo se investiga sobre el comportamiento de los métodos de fusión de imágenes más habituales para determinar cuál de ellos resulta más apropiado para su posible inclusión en una cámara digital con dos perspectivas: 1) análisis cualitativo y 2) tiempo de procesamiento.

Gracias al avance tecnológico experimentado en los últimos tiempos por los equipos de captura, así como los que se suponen en el futuro, cabe o cabrá sin duda la posibilidad de captar varias imágenes con diferentes focales y por tanto con diferentes profundidades de campo con diferencias de tiempos que podrían considerarse simultáneas. Por otra parte también dentro de los avances tecnológicos, hoy en día, los dispositivos incluyen programas software con implementaciones en hardware con distintas finalidades. Pensemos por ejemplo en las cámaras que incorporan detectores automáticos de caras. Pues bien, es evidente que nada ha de impedir la incorporación de métodos de fusión de imágenes, por supuesto en color. Esto permitiría los usuarios o profesionales obtener imágenes con diferentes enfoques de suerte que al final se obtenga todo el escenario enfocado por partes, es decir cada imagen contendrá unas partes enfocadas y otras desenfocadas. Gracias a los métodos de fusión será

posible realizar la fusión de todas esas imágenes para obtener una imagen totalmente nítida de toda la escena. El cambio de la distancia focal para obtener las diferentes profundidades de campo se puede realizar recorriendo todo el rango posible de focales. En estos barridos existen ya métodos automáticos para determinar la calidad del enfoque de los objetos en la escena, lo cual combinado con los procedimientos de fusión permitirán la obtención de imágenes nítidas de toda la escena observada y de alta calidad.

La automatización del proceso permitirá sin duda la incorporación de estos mecanismos de captura de imágenes en sistemas robotizados, entre los que se encuentran los vehículos no tripulados de distinta índole, tales como aéreos, terrestres o marinos, actualmente en fase de expansión. Gracias a lo cual será posible obtener imágenes de relativamente buena calidad.

El trabajo que se presenta se orienta hacia la investigación de métodos de fusión, dejando para posteriores investigaciones la integración de los mismos con procedimientos de enfoque automático (Maik, Cho, Shin, & Paik, 2007) así como el registro o alineación de las imágenes, que será necesario si no se utiliza un trípode durante la realización de las tomas (Zitová & Flusser, 2003).

Por tanto, los objetivos que se plantean en este trabajo son los que se mencionan a continuación:

- 1) Aprender a manejar referencias bibliográficas, así como la forma de abordar las investigaciones.
- 2) Identificar métodos de fusión de imágenes existentes en la literatura.
- 3) Determinar los métodos y procedimientos más apropiados tanto desde el punto de vista cualitativo como de tiempos de cómputo.
- 4) Analizar las ventajas e inconvenientes de cada uno de ellos, así como su problemática.
- 5) Establecer las conclusiones más relevantes obtenidas tras el desarrollo de la investigación.

1.3. Metodología

Para llevar a cabo la investigación, el planteamiento seguido respecto de la metodología es la que se describe a continuación.

La bibliografía analizada es abundante en métodos de fusión en imágenes de grises y bajo unas determinadas condiciones de resolución, principalmente espacial. Por tanto, los primeros estudios sobre fusión se centran en este tipo de imágenes. Como continuación natural el siguiente paso consiste en la extensión y adaptación de los métodos de fusión a imágenes en color, que son las que realmente proporcionan mayoritariamente los sistemas de captura, desde las webcam hasta los sistemas más sofisticados. Tras la mencionada extensión se realiza el correspondiente análisis cualitativo y en tiempo de ejecución.

Los resultados obtenidos tanto en grises como en colores permiten extraer las conclusiones oportunas respecto de los métodos de fusión analizados.

1.4. Organización del trabajo

El presente trabajo se organiza de forma que en el capítulo dos se estudian los métodos de fusión más relevantes encontrados en la literatura, en este caso orientados al procesamiento de imágenes de grises. En el capítulo tres se analiza la extensión de los métodos de fusión para imágenes en color. En el capítulo cuatro se aborda la problemática relacionada con la dimensionalidad de las imágenes a fusionar. En el capítulo cinco se muestran los resultados proporcionados por los diferentes métodos de fusión tanto en grises como en color, realizando el correspondiente análisis al respecto. Finalmente, en el capítulo seis se presentan las conclusiones pertinentes.

Capítulo 2

Revisión de métodos de fusión

2.1. Introducción

A continuación realizamos una revisión bibliográfica sobre los métodos de fusión de imágenes presentes en la bibliografía de visión por computador relativa al tema que nos ocupa, utilizando como base comparativas como (Sadjadi, 2005) y los métodos implementados en (Rockinger, 1999). Aunque nuestro objetivo consiste en unir dos o más imágenes obtenidas con distinta distancia focal en una única imagen en la que todas las zonas tengan el mejor enfoque posible, algunos de los métodos analizados se han presentado con otros fines como unir la información obtenida con un sensor de infrarrojos con una imagen del espectro visible para detectar objetivos militares (Ramac, Uner, & Varshney, 1998), combinar información procedente de distintos tipos de sensores digitales, como de luz visible e infrarroja, o unir distintos tipos de imágenes médicas para lograr una mayor resolución (Li, Manjunath, & Mitra, 1994).

En principio puede parecer que estos objetivos son distintos al que buscamos aquí, sin embargo podemos considerar cualquier método de fusión de imágenes como aquél que a partir de un conjunto de imágenes dadas nos permite obtener una nueva imagen que contiene la máxima información posible de todas ellas. Si tenemos en cuenta este rasgo común, algunos métodos que en principio parecían tener un objetivo distinto al nuestro ahora caben en el estudio que realizamos a continuación.

Los métodos se han probado con distintas imágenes, de las cuales la Figura 1 es un ejemplo representativo que utilizaremos como base para el estudio que se presenta en este trabajo. Como se puede apreciar, estas dos fotografías tienen distinto foco y la distancia entre el objeto en primer plano y el fondo hace que la parte desenfocada sea prácticamente irreconocible. Son, por tanto, dos imágenes idóneas para ilustrar los comportamientos de los métodos de fusión a estudiar, que se describen seguidamente en las secciones 2.4 a 2.10.



Figura 1: Dos imágenes con distinto foco

2.2. Media aritmética

Cuando tratamos con métodos de fusión de imágenes, el primer método que surge es el cálculo de la media de las dos imágenes de entrada, realizado píxel a píxel. Es decir, partiendo de dos imágenes I_1 e I_2 , la imagen resultante I tras la fusión de las dadas se define como:

$$I(x, y) = \frac{I_1(x, y) + I_2(x, y)}{2} \quad (1)$$

En esta ecuación presentamos la notación que utilizaremos en el resto del trabajo. Consideramos que toda imagen en escala de grises I es una matriz bidimensional, en la que cada elemento $I(x, y)$ es un píxel cuyo rango de valores varía desde 0 para el negro hasta el 255 para el blanco.

Si aplicamos este método a las imágenes con las que estamos mostrando los comportamientos de los métodos, el resultado es el que se puede ver en la Figura 2.



Figura 2: Fusión por media aritmética

2.3. Media ponderada

El principal problema del uso de la media aritmética como método de fusión es que en la imagen resultante se le da el mismo peso a las zonas con información nítida que a las zonas borrosas, dando lugar a que no se tengan en cuenta los detalles de la imagen de una forma adecuada. Este resultado se podría mejorar expresando cada píxel como una combinación lineal de los píxeles de las imágenes de partida, asignando más peso al píxel que tenga un nivel de detalle mayor, esto es, al píxel más nítido.

Para ello se definen dos matrices de pesos w_1 y w_2 , tales que para todos sus elementos se cumple $0 \leq w_1, w_2 \leq 1$ y $w_1(x, y) + w_2(x, y) = 1$. De esta forma, la imagen resultante se puede obtener como una media ponderada por medio de la siguiente ecuación:

$$I(x, y) = w_1(x, y)I_1(x, y) + w_2(x, y)I_2(x, y) \quad (2)$$

El problema que surge ahora es que para aplicar este método es necesario asignar un valor a cada uno de los pesos. Una posible solución sería utilizar la información procedente de los bordes en los objetos de las imágenes (Haerberli, 1994), obtenida, por ejemplo, aplicando un filtro de paso alto a las imágenes bajo fusión. Cuando hablamos de altas frecuencias en una imagen nos referimos a aquellos componentes que cambian rápidamente, este es el caso precisamente de los bordes de la imagen. Las bajas frecuencias de una imagen serían, por su parte, aquellos píxeles con poca variación en su intensidad respecto a su entorno, las zonas homogéneas desde el punto de vista de la intensidad de la imagen. En el Anexo I se ha incluido información sobre estos conceptos de altas y bajas frecuencias.

Un filtro de paso alto se puede definir de diversas maneras. En nuestro caso utilizaremos un núcleo que tiene propiedades de tipo paso alto, característico en tratamiento de imágenes, dicho núcleo en el dominio espacial permite realizar una convolución sobre cada imagen a ser fusionada de la siguiente forma:

$$H(x, y) = |h * I(x, y)| \quad (3)$$

donde la máscara h elegida es exactamente de tipo Gaussiano procedente de la aplicación del concepto segunda derivada, se define como:

$$h = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \quad (4)$$

Así, si H_1 contiene la información de bordes de la imagen I_1 y H_2 la de la imagen I_2 , en ambos casos por aplicación de la máscara definida previamente, podemos definir los pesos como:

$$w_1(x, y) = \frac{H_1(x, y)}{H_1(x, y) + H_2(x, y)} \quad (5)$$

$$w_2(x, y) = 1 - w_1(x, y) \quad (6)$$

En aquellos puntos en los que el denominador de (5) sea nulo, es decir, no haya bordes en ninguna de las dos imágenes, se asignaría a w_1 el valor 0,5. En la Figura 3 podemos ver el valor que toman estos pesos para cada una de las dos imágenes: las zonas claras indican bordes, las zonas oscuras indican zonas homogéneas.

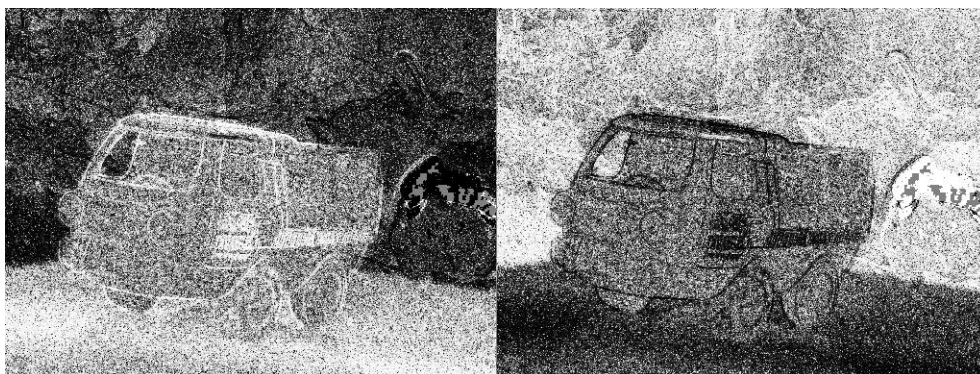


Figura 3: Pesos correspondientes a las imágenes

Un problema del uso de este tipo de fusión, como se señala en (Mertens, Kautz, & Van Reeth, 2007), es que si hay regiones en las que los pesos cambian rápidamente, el resultado puede no ser satisfactorio ya que la imagen resultante tendrá una especie de “costuras”, técnicamente son fenómenos conocidos también en la literatura especializada como artefactos. Podemos apreciar estas “costuras” en la Figura 4, que representa la imagen resultante de la fusión por media ponderada, en ella podemos ver cómo el techo del coche de juguete está enfocado, justo encima del mismo aparece una franja en la que está enfocado el fondo y sobre la mencionada franja aparece de nuevo el techo del coche a modo de repetición. El mismo efecto se puede apreciar a la izquierda del tronco de la derecha, por tanto se trata de un efecto característico en este tipo de situaciones



Figura 4: Fusión por media ponderada

2.4. Pirámide de diferencia de paso bajo

En el método anterior hemos introducido un concepto nuevo consistente en usar un filtro paso alto para saber dónde están los detalles de la imagen. El método de la pirámide de diferencia de paso bajo (DOLP), también conocido como pirámide Laplaciana, utiliza una aproximación similar.

Se basa en aplicar recursivamente un filtro paso bajo y una reducción de tamaño a la imagen para obtener una serie de imágenes cada vez con menos detalle y menor dimensión a partir de dichas operaciones (coeficientes de aproximación o pirámide gaussiana) y guardar la diferencia entre esas imágenes y las originales (coeficientes de detalle o pirámide Laplaciana) (Burt & Kolczynski, 1993). Si aplicamos esta descomposición en pirámide sobre una imagen volviéndola a reconstruir posteriormente, obtenemos la imagen original sin pérdida (a excepción de posibles errores de redondeo).

Para una imagen I , definimos el nivel 0 de su pirámide gaussiana como la copia píxel a píxel de la imagen original, mientras que su nivel k -ésimo, con $k \geq 1$, tendría la forma

$$G_k(x, y) = [w * G_{k-1}(x, y)]_{\downarrow 2} \quad (7)$$

La notación $\downarrow 2$ se refiere a eliminar la mitad de filas y columnas; posteriormente veremos la notación análoga $\uparrow 2$, que representa el hecho de duplicar las filas y columnas. La matriz w es el núcleo de la convolución o máscara de desenfoque y viene dada por:

$$w = \frac{1}{256} \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix} \quad (8)$$

El nivel k de la pirámide Laplaciana, por su parte, se define del siguiente modo:

$$\tilde{L}_k(x, y) = G_k(x, y) - 4w * [G_{k+1}(x, y)]_{\uparrow 2} \quad (9)$$

La Figura 5 muestra las pirámides gaussiana (a la izquierda) y Laplaciana (a la derecha) correspondientes a la primera fotografía de la Figura 1 para un total de cinco niveles.

Para reconstruir la imagen a partir de la pirámide, se aplica la siguiente ecuación:

$$\hat{G}_k(x, y) = \tilde{L}_k(x, y) + 4w * [\hat{G}_{k+1}(x, y)]_{\uparrow 2} \quad (10)$$

donde \hat{G}_0 sería la reconstrucción de la imagen original I .

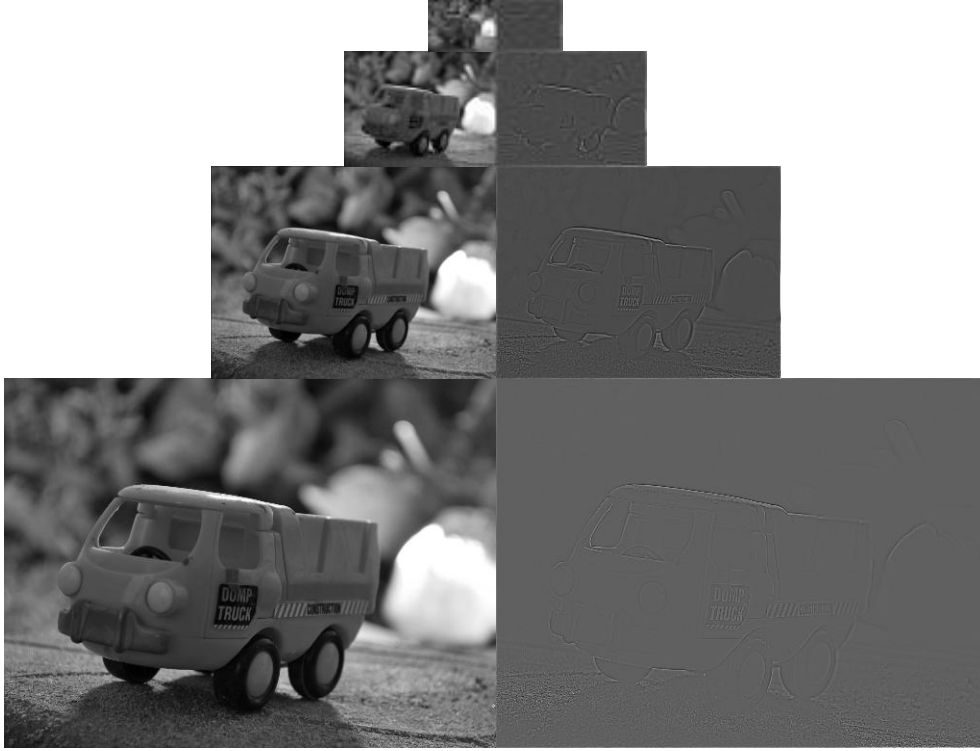


Figura 5: Pirámides gaussiana y Laplaciana

Si en lugar de aplicar este método sobre una única imagen lo hacemos sobre dos o más, podremos combinar los coeficientes de aproximación G_k y de detalle \tilde{L}_k de distintas maneras, dando lugar a la fusión de las imágenes originales.

El modo de combinar ambos tipos de coeficientes no tiene por qué ser el mismo: en primer lugar los valores de los coeficientes de aproximación se situarán en el rango de 0 a 255, al igual que los valores de la imagen original, mientras que los coeficientes de detalle tendrán valores próximos a cero. Además, los coeficientes de detalle se deben fusionar en todos los niveles, mientras que los coeficientes de aproximación sólo tienen que combinarse en la última capa. La solución propuesta en (Burt, van der Wal, Kolczynski, & Hingorani, 1993) es fusionar estos residuos de los coeficientes de aproximación utilizando la media aritmética y combinar los coeficientes de detalle escogiendo el que tenga un mayor valor absoluto, lo que se conoce como selección por *medida de actividad* basada en coeficientes (Zhang & Blum, 1999), del siguiente modo:

$$G_k^C(x, y) = \frac{G_k^A(x, y) + G_k^B(x, y)}{2} \quad (11)$$

$$\tilde{L}_k^C(x, y) = \begin{cases} \tilde{L}_k^A(x, y) & \text{si } |\tilde{L}_k^A(x, y)| > |\tilde{L}_k^B(x, y)| \\ \tilde{L}_k^B(x, y) & \text{e. o. c.} \end{cases} \quad (12)$$

En la Figura 6 podemos ver el resultado de fusionar las dos imágenes de la Figura 1 utilizando el método descrito previamente. En este caso en la descomposición se ha llegado a un máximo de 8 niveles.



Figura 6: Fusión por pirámide de diferencia de paso bajo

2.5. Pirámide Laplaciana FSD

En la sección anterior vimos que necesitábamos filtrar la imagen original y reducirla para obtener la gaussiana y a continuación expandirla, volver a filtrarla y realizar una operación de resta para obtener la Laplaciana. Esta reducción y expansión determina que nos refiramos a esta pirámide como Laplaciana RE.

La pirámide Laplaciana FSD (*filter, subtract, decimate* o filtrado, resta y reducción del tamaño) propuesta en (Anderson, 1984) es una variación específicamente pensada para ser computacionalmente más eficiente, en la que definimos el nivel k de la pirámide Laplaciana como

$$L_k(x, y) = G_k(x, y) - w * G_k(x, y) \quad (13)$$

Como podemos ver ya no es necesaria la expansión de la ecuación (9). Sin embargo, como se apunta en (Anderson, 1987), ya no es posible usar (10) para reconstruir la imagen, pues se producirían pérdidas en el detalle, por lo que es necesario utilizar una reconstrucción diferente tal como la propuesta en la siguiente ecuación:

$$\hat{G}_k(x, y) = L_k(x, y) + w * \left(L_k(x, y) + [4\hat{G}_{k+1}(x, y)]_{\uparrow 2} \right) \quad (14)$$

En la Figura 7 podemos apreciar la descomposición piramidal aplicada sobre la imagen de la Figura 1.

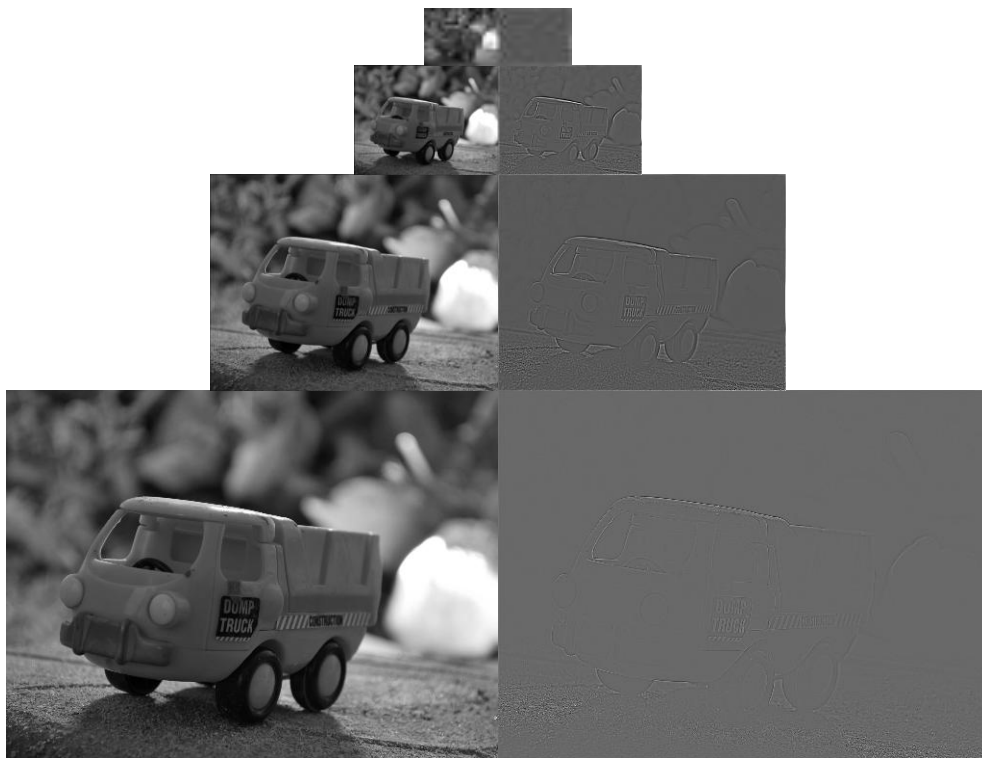


Figura 7: Pirámides gaussiana y Laplaciana FSD

Realizando la fusión igual que con la pirámide DOLP, es decir, haciendo una media de las pirámides gaussianas y tomando el mayor valor absoluto de las Laplacianas, obtenemos un resultado como el mostrado en la Figura 8.



Figura 8: Fusión por pirámide Laplaciana FSD

2.6. Pirámide de razón de paso bajo

El método de la razón de paso bajo o ROLP (Toet, 1989), (Toet, 1990), también conocido como pirámide del contraste (Toet, van Ruyven, & Valetton, 1989) es otra modificación del método DOLP. En este caso el objetivo no es tener un algoritmo más eficiente sino imitar el comportamiento del ojo humano. En efecto, el ojo es más sensible a variaciones en el contraste de una imagen que a variaciones en el valor absoluto, por tanto en este método en lugar de utilizarse diferencias (restas) se utilizan razones (divisiones).

Definiendo la pirámide gaussiana igual que en (7), el nivel k de la pirámide de razones sería:

$$R_k(x, y) = G_k(x, y) / 4w * [G_{k+1}(x, y)]_{\uparrow 2} \quad (15)$$

En la Figura 9 podemos ver una pirámide ROLP generada usando esta ecuación.

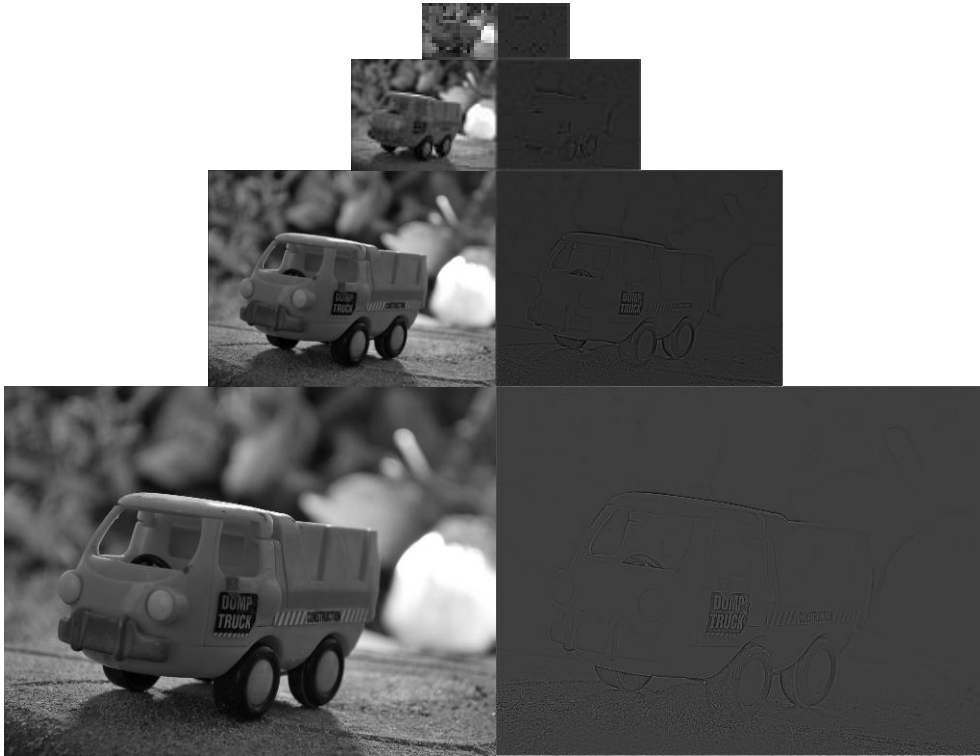


Figura 9: Pirámide gaussiana y de razón de paso bajo

Para reconstruir la imagen se aplicaría la siguiente expresión:

$$\hat{G}_k(x, y) = R_k(x, y) \cdot 4w * [\hat{G}_{k+1}(x, y)]_{\uparrow 2} \quad (16)$$

A la hora de combinar los coeficientes de detalle R_k^A y R_k^B para dos imágenes, en lugar de tomar el máximo directamente, como hacíamos en el caso de las dos pirámides anteriores, se aplicará el criterio proporcionado por la siguiente ecuación:

$$R_k^C(x, y) = \begin{cases} R_k^A(x, y) & \text{si } |R_k^A(x, y) - 1| > |R_k^B(x, y) - 1| \\ R_k^B(x, y) & \text{e. o. c.} \end{cases} \quad (17)$$

Aplicando este método sobre las imágenes que venimos utilizando obtenemos el resultado mostrado en la imagen de la Figura 10.



Figura 10: Fusión por pirámide ROLP

2.7. Pirámide del gradiente

La pirámide del gradiente (Burt & Kolczynski, 1993), (Burt, 1992) también surge como una variación del método DOLP, en la que ahora no se guarda sólo la pirámide gaussiana sino cuatro pirámides distintas, cada una con una orientación diferente (horizontal, vertical y las dos diagonales).

El nivel k con orientación l de la pirámide del gradiente se define como:

$$D_{k,l}(x, y) = d_l * [G_k(x, y) + \hat{w}G_k(x, y)] \quad (18)$$

Donde G_k es el nivel k de la pirámide gaussiana obtenida con (7), d_l es el filtro del gradiente para la orientación l y \hat{w} es el núcleo de desenfoque gaussiano. Estos filtros se definen del siguiente modo:

$$d_1 = [1 \quad -1] \quad (19)$$

$$d_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad (20)$$

$$d_3 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad (21)$$

$$d_4 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \quad (22)$$

$$\dot{w} = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad (23)$$

En la Figura 11 podemos ver las cuatro orientaciones de la pirámide del gradiente para un único nivel de descomposición. De izquierda a derecha: horizontal, diagonal secundaria, vertical, diagonal principal.



Figura 11: Orientaciones de la pirámide del gradiente

Para reconstruir la imagen original a partir de la pirámide del gradiente podemos obtener la pirámide Laplaciana FSD del siguiente modo:

$$L_k(x, y) = \sum_{l=1}^4 \left(-\frac{1}{8} d_l * D_{k,l}(x, y) \right) \quad (24)$$

En lugar de aplicar directamente la reconstrucción (14), se propone una aproximación que permite obtener la pirámide Laplaciana RE a partir de la pirámide FSD del siguiente modo:

$$\tilde{L}_k(x, y) \approx [1 + w] * L_k(x, y) \quad (25)$$

Con lo que finalmente la pirámide Laplaciana que usaremos será

$$\tilde{L}_k(x, y) = [1 + w] * \sum_{l=1}^4 \left(-\frac{1}{8} d_l * D_{k,l}(x, y) \right) \quad (26)$$

A partir de la pirámide Laplaciana obtenida con (26) se puede recuperar la imagen original usando (10). Es importante señalar que cuando hablamos del núcleo $1 + w$, lo que hacemos es sumar 1 a la posición central de la máscara w y no a todos los elementos. Es decir, a partir de (8) tenemos:

$$1 + w = \frac{1}{256} \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 292 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix} \quad (27)$$

Tras aplicar una fusión por pirámide del gradiente obtenemos la Figura 12.



Figura 12: Fusión por pirámide del gradiente

2.8. Pirámide morfológica

Todos los métodos que hemos visto hasta ahora se basaban principalmente en la descomposición de imágenes en bajas y altas frecuencias. La pirámide morfológica (Ramac, Uner, & Varshney, 1998), (Heijmans & Goutsias, 2005), sin embargo, aplica una idea diferente y sigue un criterio distinto.

Para obtener los coeficientes de aproximación de una imagen, aplica sobre ella los operadores morfológicos de apertura y cierre del siguiente modo:

$$X_k(x, y) = [(X_{k-1}(x, y) \circ B) \cdot B]_{\downarrow 2} \quad (28)$$

A continuación, los coeficientes de detalle, que permitirán reconstruir la imagen, se obtienen del siguiente modo:

$$Y_k(x, y) = X_k(x, y) - ([X_{k+1}(x, y)]_{\uparrow 2} \oplus B) \quad (29)$$

En ambos casos, el elemento estructural B es el mismo, elegido de antemano. La Figura 13 muestra el resultado de aplicar el criterio de la pirámide morfológica a la imagen del ejemplo.

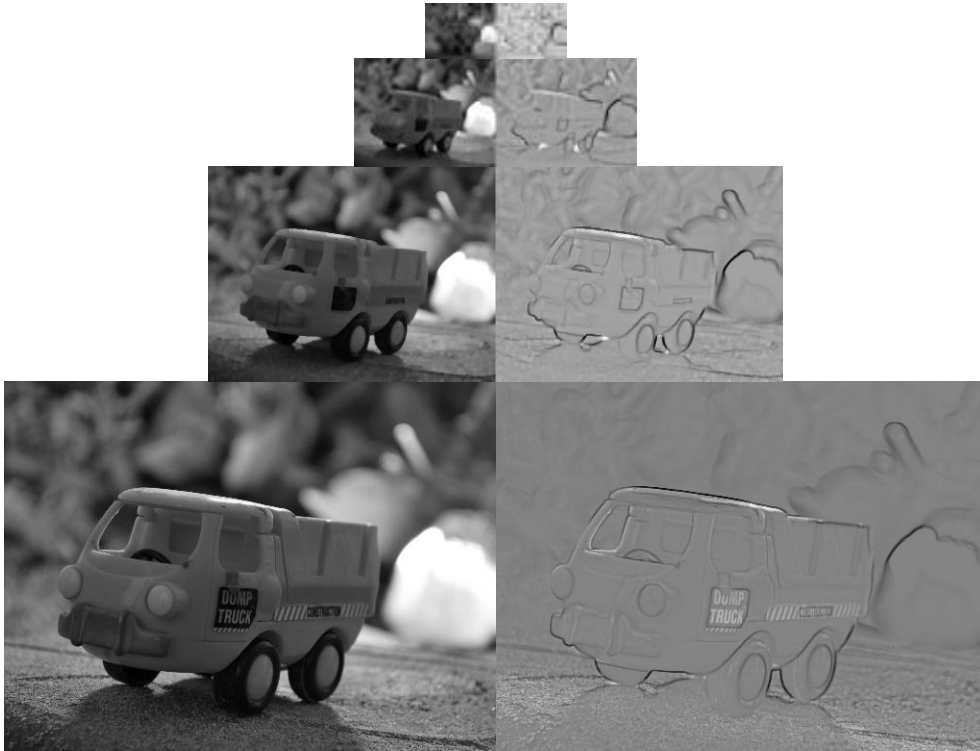


Figura 13: Pirámide morfológica

Para reconstruir la imagen tras su descomposición se aplicará la ecuación

$$X_k(x, y) = Y_k(x, y) + ([X_{k+1}(x, y)]_{\uparrow 2} \oplus B) \quad (30)$$

En la Figura 14 podemos ver el resultado de esta fusión.



Figura 14: Fusión por pirámide morfológica

2.9. Transformada de wavelets

La transformada discreta de *wavelets* (Mallat, 1989), (Pajares & de la Cruz, 2004), también conocida como DWT por sus iniciales en inglés, se basa en una teoría similar a la de la pirámide del gradiente, salvo que en este caso no es necesario almacenar la pirámide gaussiana y es suficiente con las cuatro pirámides correspondientes a los coeficientes de aproximación (bajas frecuencias de la imagen) y los coeficientes de detalle horizontales, verticales y diagonales (altas frecuencias en cada uno de los ejes o en ambos).

En líneas generales, definimos el nivel k de las pirámides de wavelets como:

$$LL_k(x, y) = [w_L^T * L_k(x, y)]_{1 \downarrow 2} \quad (31)$$

$$LH_k(x, y) = [w_H^T * L_k(x, y)]_{1 \downarrow 2} \quad (32)$$

$$HL_k(x, y) = [w_L^T * H_k(x, y)]_{1 \downarrow 2} \quad (33)$$

$$HH_k(x, y) = [w_L^T * H_k(x, y)]_{1 \downarrow 2} \quad (34)$$

Donde la notación $1 \downarrow 2$ se refiere a eliminar la mitad de las filas de la imagen (dejando intacto el número de columnas) y L y H son pasos intermedios, obtenidos de la imagen original I del siguiente modo:

$$L_k(x, y) = [w_L * I_k(x, y)]_{2 \downarrow 1} \quad (35)$$

$$H_k(x, y) = [w_H * I_k(x, y)]_{2 \downarrow 1} \quad (36)$$

En todos los casos w_L y w_H son los núcleos de paso bajo y paso alto respectivamente. En el caso de la transformada de wavelets se pueden utilizar cualquiera de los definidos para las diferentes familias de wavelets (Amolins, Zhang, & Dare, 2007), siendo relativamente habitual utilizar uno de los más sencillos como son los de Haar, que vienen dados por:

$$w_L = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \end{bmatrix} \quad (37)$$

$$w_H = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \end{bmatrix} \quad (38)$$

La pirámide se forma aplicando esta descomposición de modo recursivo en la subimagen LL , es decir, sobre los coeficientes de aproximación. La Figura 15 muestra un nivel de descomposición de la pirámide DWT: las subimágenes superiores se corresponden con LL y LH mientras que las inferiores con HL y HH .

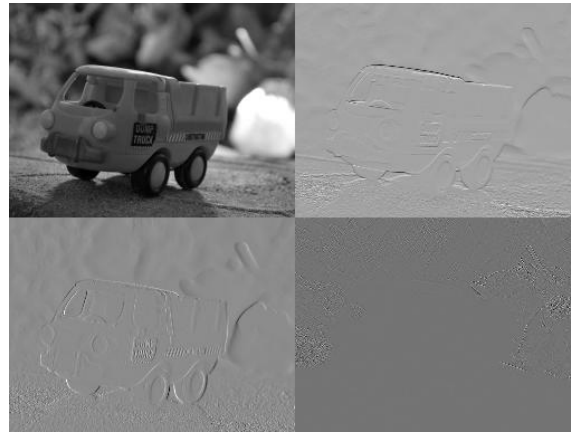


Figura 15: Un nivel de la pirámide DWT

Para realizar la reconstrucción se aplican las mismas transformaciones sobre las cuatro subimágenes del siguiente modo:

$$LL_k^{-1}(x, y) = w_{L^{-1}}^T * [w_{L^{-1}} * LL_k(x, y)_{1\uparrow 2}]_{2\uparrow 1} \quad (39)$$

$$LH_k^{-1}(x, y) = w_{H^{-1}}^T * [w_{L^{-1}} * LH_k(x, y)_{1\uparrow 2}]_{2\uparrow 1} \quad (40)$$

$$HL_k^{-1}(x, y) = w_{L^{-1}}^T * [w_{H^{-1}} * HL_k(x, y)_{1\uparrow 2}]_{2\uparrow 1} \quad (41)$$

$$HH_k^{-1}(x, y) = w_{H^{-1}}^T * [w_{H^{-1}} * HH_k(x, y)_{1\uparrow 2}]_{2\uparrow 1} \quad (42)$$

donde los núcleos $w_{L^{-1}}$ y $w_{H^{-1}}$ se definen como:

$$w_{L^{-1}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \end{bmatrix} \quad (43)$$

$$w_{H^{-1}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \end{bmatrix} \quad (44)$$

Por último, la imagen reconstruida se obtiene a partir de la siguiente expresión:

$$\hat{I} = LL_k^{-1}(x, y) + LH_k^{-1}(x, y) + HL_k^{-1}(x, y) + HH_k^{-1}(x, y) \quad (45)$$

En la Figura 16 se puede apreciar el resultado de la fusión realizada por aplicación de este método:



Figura 16: Fusión por pirámide DWT

2.10. Transformada de wavelets invariante a traslaciones

La transformada de *wavelets* invariante a traslaciones (Rockinger, 1997), o SIDWT, surgió como solución a un problema característico de todos los métodos de descomposición en pirámide que se hace especialmente acusado en el caso de la pirámide DWT: las transformadas de dos imágenes que sean prácticamente iguales, pero desplazadas unos píxeles, son muy diferentes (Beylkin, 1992).

Para evitar esto la pirámide SIDWT propone modificar características de la pirámide DWT. En primer lugar se eliminan todos los redimensionados y en segundo lugar los filtros w_L y w_H serán distintos para cada nivel de la pirámide, añadiendo cada vez un cierto número de ceros entre sus coeficientes para evitar las distorsiones de fase.

La Figura 17 muestra un ejemplo de esta fusión.



Figura 17: Fusión por pirámide SIDWT

2.11. Otros métodos de combinación de coeficientes

En todos los casos que hemos visto la combinación de los coeficientes correspondientes a bajas frecuencias se hacía realizando directamente la media, como en el caso de (11), y la de altas frecuencias escogiendo la que tuviera un mayor valor absoluto como en (12).

Estos métodos de combinación no son los únicos que podemos usar. En el caso de los coeficientes de detalle, por ejemplo, en lugar de tomar directamente el máximo podemos tener en cuenta el entorno de un píxel para decidir cuál usar. En el caso de la pirámide DOLP se haría del siguiente modo:

$$\tilde{L}_k^C(x, y) = \begin{cases} \tilde{L}_k^A(x, y) & \text{si } w * |\tilde{L}_k^A(x, y)| > w * |\tilde{L}_k^B(x, y)| \\ \tilde{L}_k^B(x, y) & \text{e. o. c.} \end{cases} \quad (46)$$

Nótese que este método es equivalente al de tomar directamente el máximo si tomamos una ventana de dimensión 1×1 , es decir $w = [1]$. Este tipo de métodos, en el que tomamos directamente el valor de una de las dos pirámides, se conoce como de selección; si sólo tenemos en cuenta los coeficientes por separado hablaremos de medidas de actividad basadas en coeficientes y si tenemos en cuenta el entorno hablaremos de medidas de actividad basadas en ventana.

Además de métodos de selección también podemos combinar coeficientes usando diferentes tipos de promediado. Es decir, mezclaremos el valor de dos coeficientes en lugar de coger uno “tal cual” como hacíamos en la selección.

El primer método de promediado que veremos será la media ponderada general. Hasta ahora habíamos usado la media aritmética con los coeficientes de aproximación. Ahora usaremos la misma base que la media ponderada, que vimos en la sección 2.3, para combinar estos coeficientes del siguiente modo:

$$G_k^C(x, y) = w_1(x, y)G_k^A(x, y) + w_2(x, y)G_k^B(x, y) \quad (47)$$

Aquí los pesos se calculan de un modo diferente al de (5) y (6). En concreto, para la media ponderada general que usaremos, en primer lugar se calcula una medida de correspondencia entre las dos pirámides de coeficientes (Adelson, Anderson, Bergen, Burt, & Ogden, 1984) del siguiente modo:

$$M_{XY}(x, y) = \frac{w * G_k^A(x, y)G_k^B(x, y)}{(w * |G_k^A(x, y)|)^2 + (w * |G_k^B(x, y)|)^2} \quad (48)$$

Para calcular los pesos a partir de M_{XY} , en primer lugar se establece un valor de umbral α . En aquellos puntos en los que $w * |G_k^A(x, y)| < w * |G_k^B(x, y)|$ tenemos que:

$$w_1(x, y) = \begin{cases} \frac{1}{2} - \frac{1}{2} \left(\frac{1 - M_{XY}(x, y)}{1 - \alpha} \right) & \text{si } M_{XY}(x, y) \geq \alpha \\ 0 & \text{e. o. c.} \end{cases} \quad (49)$$

$$w_2(x, y) = 1 - w_1(x, y) \quad (50)$$

En el resto de puntos los pesos intercambiarían sus subíndices. Aquí hemos usado la pirámide DOLP pero este método puede usarse para cualquier pirámide en la que distingamos entre bajas y altas frecuencias.

Para los coeficientes de detalle utilizaremos un promediado diferente: la media ponderada adaptativa. Este método se basa en otorgar un mayor valor a aquellos puntos en los que haya una mayor diferencia respecto a los píxeles del entorno, de modo que tenemos la siguiente expresión

$$W_1 = (|\tilde{L}_k^A(x, y) - w * \tilde{L}_k^A(x, y)|)^a \quad (51)$$

donde α es un parámetro que permite modificar la distribución de los pesos.

Para obtener las matrices de pesos podemos proceder del siguiente modo:

$$w_1(x, y) = \frac{W_1(x, y)}{W_1(x, y) + W_2(x, y)} \quad (52)$$

Por su parte w_2 se definiría igual que en (50).

Por último estudiaremos un método de combinación de coeficientes basado en el gradiente global de la imagen (Santos, Pajares, Portela, & de la Cruz, 2003). En este caso veremos este método aplicado en la pirámide DWT aunque su uso se puede extrapolar al resto de técnicas de fusión.

En primer lugar los coeficientes de aproximación se escogerán del siguiente modo:

$$LL_k^C(x, y) = \begin{cases} LL_k^A(x, y) & \text{si } |h * LL_k^A(x, y)| > |h * LL_k^B(x, y)| \\ LL_k^B(x, y) & \text{e. o. c.} \end{cases} \quad (53)$$

donde h es la máscara definida en la ecuación (4).

Para los coeficientes de detalle el procedimiento es ligeramente diferente:

$$HH_k^C(x, y) = \begin{cases} HH_k^A(x, y) & \text{si } g_k^A(x, y) > g_k^B(x, y) \\ HH_k^B(x, y) & \text{e. o. c.} \end{cases} \quad (54)$$

donde definimos g_k como

$$g_k(x, y) = |w * LH_k^A(x, y) + w * HL_k^A(x, y) + w * HH_k^A(x, y)| \quad (55)$$

El resto de coeficientes de detalle se calcularía del mismo modo.

Capítulo 3

Fusión en imágenes de color

La Figura 18 muestra dos imágenes en color. Con los métodos que hemos visto en el capítulo anterior no podríamos fusionar estas dos imágenes, ya que hasta ahora en los métodos descritos considerábamos las imágenes como matrices bidimensionales en las que cada píxel tenía un valor entre 0 y 255.



Figura 18: Dos imágenes a color con distintos grados de desenfoque

En las imágenes en color no podemos seguir tratando los píxeles del mismo modo. Por tanto antes de ver cómo podemos modificar los métodos para fusionar este tipo de imágenes, vamos a estudiar los modelos de representación del color o espacios de color desde el punto de vista de la fusión.

3.1. Espacios de color

En una imagen en una pantalla, ya sea un monitor de ordenador, una televisión o una imagen proyectada, cada píxel de color posee tres componentes: rojo, verde y azul. Estos son los colores primarios de la luz y cualquier color que podamos percibir se puede formar a partir de ellos. Dado que el modo más habitual que tiene un ordenador de mostrar imágenes es a través de su pantalla, éste es el método más habitual que se suele emplear para almacenar las imágenes, esto es: sus componentes rojo, verde y azul, o RGB por sus iniciales en inglés. La Figura 19 muestra las respectivas componentes roja, verde y azul de una imagen obtenidas a partir de la imagen de la Figura 18.



Figura 19: Componentes RGB de una imagen

Bien es cierto que este método de almacenar el color no es el único. Sin ir más lejos, a la hora de imprimir una imagen los colores primarios son otros: cian, amarillo y magenta (CYM en inglés). El ordenador debe realizar una conversión de uno a otro formato para que la impresora pueda representar fielmente la imagen. Sin embargo, así como la pantalla podía mostrar cualquier color “por el mismo precio” si queremos imprimir algo de color negro usando este método la impresora consumirá tinta de tres colores. Como la tinta negra es habitualmente mucho más barata que las de color, las impresoras normalmente tienen cuatro cartuchos: cian, amarillo, magenta y negro (cuatricromía conocida habitualmente como CMYK).

Además de los modelos de color mencionados existen otros varios. Sin necesidad de entrar en detalles, en el caso de imágenes en vídeo se utilizan espacios de color como YUV (en el sistema de televisión PAL), YIQ (en NTSC) y YCbCr (similar a YUV). Todos ellos comparten una división en tres canales, uno de luminancia y dos de “crominancia” o información de color. En el caso de YUV y YIQ el canal de luminancia representa la imagen original en escala de grises conservando los detalles de la misma; en el caso de YCbCr, más orientado al video digital, los canales se encuentran “comprimidos” para poder añadir datos adicionales en el flujo de video (The MathWorks, 1984-2009). La Figura 20 muestra la descomposición YIQ a partir de la imagen RGB mostrada en la Figura 18.



Figura 20: Componentes YIQ de una imagen

Otros espacios de color ampliamente utilizados son HSV (tono, saturación y valor) y HSI (tono, saturación e intensidad), especialmente presentes en las interfaces de usuario de programas de edición de imagen, por la facilidad que presentan para escoger un color en concreto cambiando su tono o su saturación. En el caso de HSI el canal I correspondiente a la intensidad es una copia en escala de grises de la imagen, similar al canal Y en YUV y YIQ, donde ahora la información de color se contiene en las componentes H y S

3.2. Espacios de color para fusión de imágenes

A la hora de realizar una fusión de imágenes en color tendremos que escoger uno de los espacios de color anteriormente mencionados, o varios si queremos comparar los resultados de la fusión por medio de diferentes representaciones en color. Dado que la imagen original se encuentra directamente almacenada en formato RGB, éste será uno de los formatos que usemos. Este espacio de color tiene el problema de que la fusión debe hacerse sobre cada uno de los tres canales, ya que los tres contienen información de detalle de la imagen. Esta fusión también debería llevarse a cabo canal a canal en el caso de utilizar el modelo CYM.

En cambio mediante la representación de otros espacios de color, tales como YIQ, podríamos realizar la fusión en detalle únicamente sobre el canal de luminancia o intensidad, realizando en los otros dos canales una fusión más simple. Por ejemplo, se podrían tomar los puntos de la imagen cuyo canal de luminancia más se pareciera al fusionado, como muestra la siguiente ecuación:

$$IQ(x, y) = \begin{cases} IQ^A(x, y) & \text{si } dist_A(x, y) < dist_B(x, y) \\ IQ^B(x, y) & \text{si } dist_B(x, y) < dist_A(x, y) \\ (IQ^A(x, y) + IQ^B(x, y))/2 & \text{e. o. c.} \end{cases} \quad (56)$$

donde $dist_A(x, y) = |Y(x, y) - Y^A(x, y)|$ y análogamente con $dist_B$.

Esto supondría un ahorro de recursos si el cambio de un espacio de color a otro no fuera computacionalmente costoso. En el caso de todos los espacios aquí vistos el menos costoso desde el punto de vista computacional resulta ser el espacio YIQ, que se consigue con la multiplicación matricial de (57), por lo tanto además del modelo de color RGB usaremos también este espacio.

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (57)$$

En resumen, para realizar fusiones de imágenes en color por un lado probaremos a fusionar los canales RGB uno a uno y por otro convertiremos la imagen al espacio YIQ. Para ambos métodos analizaremos tanto el tiempo necesario para la ejecución como el error obtenido, consiguiendo así las pertinentes conclusiones sobre lo que realmente merece la pena, ya sea fusionar directamente a partir del modelo original o previa transformación del espacio de color.

Capítulo 4

Fusión en imágenes con dimensiones inadecuadas

Un problema que comparten muchos de los métodos aquí vistos es que en cada paso dividen el ancho y el alto de la imagen original por dos. Esto hace imprescindible que tanto el ancho como el alto de la imagen sean divisibles por dos tantas veces como queramos aplicar el algoritmo de fusión correspondiente.

Las imágenes de la Figura 18 se caracterizan por poseer unas dimensiones de 3872×2592 píxeles. Esto quiere decir que, como máximo, se podrán dividir por dos un total de cinco veces, por lo que no podremos pasar de ese nivel a la hora de crear las pirámides. Para resolver este problema vamos a probar tres estrategias diferentes para comprobar cuál de ellas proporciona los mejores resultados.

4.1. Usar las dimensiones originales

La primera estrategia consistirá en no reducir las imágenes. Omitir el paso de reducir el tamaño de la imagen nos daría un nuevo juego de ecuaciones para cada uno de los métodos vistos en el Capítulo 2. Por ejemplo, el método de la pirámide de diferencia de paso bajo daría lugar a las siguientes definiciones:

$$G_k(x, y) = w * G_{k-1}(x, y) \quad (58)$$

$$\tilde{L}_k(x, y) = G_k(x, y) - G_{k+1}(x, y) \quad (59)$$

$$\hat{G}_k(x, y) = \tilde{L}_k(x, y) + \hat{G}_{k+1}(x, y) \quad (60)$$

donde G_0 sería la imagen original I , y \hat{G}_0 sería su reconstrucción.

4.2. Añadir bandas negras

La segunda estrategia será añadir a las imágenes originales tantas filas y columnas negras (con valor 0) como sea necesario para poder realizar la fusión, eliminando dichas filas y columnas después de realizar la fusión. Este método es el empleado en (Rockinger, 1999).

Para que las dimensiones de las imágenes de la Figura 18 fueran divisibles por dos hasta ocho veces sería necesario que sus dimensiones fueran 4096×2816 . Estos píxeles que faltan los obtenemos añadiendo las bandas negras necesarias, de modo que las imágenes que usaríamos para llevar a cabo la fusión serían las ampliadas que se pueden apreciar en sendas imágenes de la Figura 21.



Figura 21: Imágenes con bandas negras

4.3. Redimensionar las imágenes

La última estrategia será redimensionar las imágenes de partida usando dos algoritmos bien conocidos, como son a) re-escalado por vecino más próximo y b) interpolación bicúbica. Esta operación de re-escalado tiene como finalidad ampliar la imagen hasta la dimensión que sea necesaria para poder realizar la fusión, devolviendo la imagen resultante al tamaño original al finalizar el proceso de fusión.

La Figura 22 muestra el resultado de la interpolación por vecino más próximo sobre las imágenes de la Figura 18 con el fin de conseguir el tamaño adecuado para realizar la fusión. En este caso 4096×2816 , como vimos en la sección anterior.



Figura 22: Imágenes interpoladas por vecino más próximo

Capítulo 5

Resultados

5.1. Resultados de los métodos de fusión

Hemos visto que muchos de los métodos de fusión del Capítulo 2 necesitaban que la imagen original estuviera en escala de grises y tanto su ancho como su alto fueran divisibles por 2 tantas veces como niveles quisiéramos que tuviera la pirámide. Por lo tanto hemos escogido un conjunto de pares de imágenes que cumplen estos requisitos. El número total de imágenes de prueba es de cuarenta, de los cuales en la Figura 23 se muestran tres a modo de ejemplos ilustrativos.



Figura 23: Algunas de las imágenes que se van a usar para las pruebas

Los pares de imágenes utilizados en el proceso de fusión son sintéticos. Es decir, cada par se ha creado a partir de una única imagen original que no presenta desenfoque alguno, procediendo posteriormente a desenfocar una parte diferente en sendas imágenes del par mediante el programa de edición fotográfica Gimp para Windows (Equipo de desarrollo de Gimp). Esto ha sido necesario para poder cuantificar el error real de los métodos, ya que de haber utilizado imágenes reales no se tendría una imagen contra la que comprobar los resultados.

Para evaluar cada uno de los métodos de fusión comprobaremos el error cuadrático medio (MSE), que se produce respecto de la imagen a partir de la cual se crearon las imágenes que se están fusionando. Este error se define como la suma de los cuadrados de los errores sobre el conjunto de píxeles dividido entre el número de píxeles de la imagen.

Igualmente comprobaremos el tiempo de procesamiento necesario para cada método. Ante la imposibilidad de implementar los métodos en hardware como en (Sims & Irvine, 2006), todas las pruebas se realizarán en software, sobre la misma máquina, con un procesador Intel Core i5 430M de doble núcleo (Intel), usando el programa Matlab 7.9.0 (The MathWorks) y bajo una plataforma Windows 7 Professional de 64 bits (Microsoft Corporation). Para conseguir tiempos fidedignos durante la realización de las pruebas no se tiene en memoria ningún otro programa y se intenta que en todos los casos las condiciones sean las mismas.

A excepción de la media y la media ponderada, los demás métodos se han estudiado para un número de niveles de la pirámide entre dos y ocho. Los resultados se recogen en tres figuras. En la Figura 24 se muestra el tiempo necesario para cada uno de los métodos en función del número de niveles de descomposición de su pirámide; la Figura 25 muestra el error cometido para cada uno de esos niveles por comparación con la correspondiente imagen original y por último la Figura 26 muestra el error cometido en función del tiempo de procesamiento. En esta última figura un método obtiene mejor resultado cuanto más abajo quede algún punto de su gráfica y será más rápido cuanto más a la izquierda esté, por lo tanto cuanto más cerca se sitúe alguno de sus puntos del origen mejor será el método.

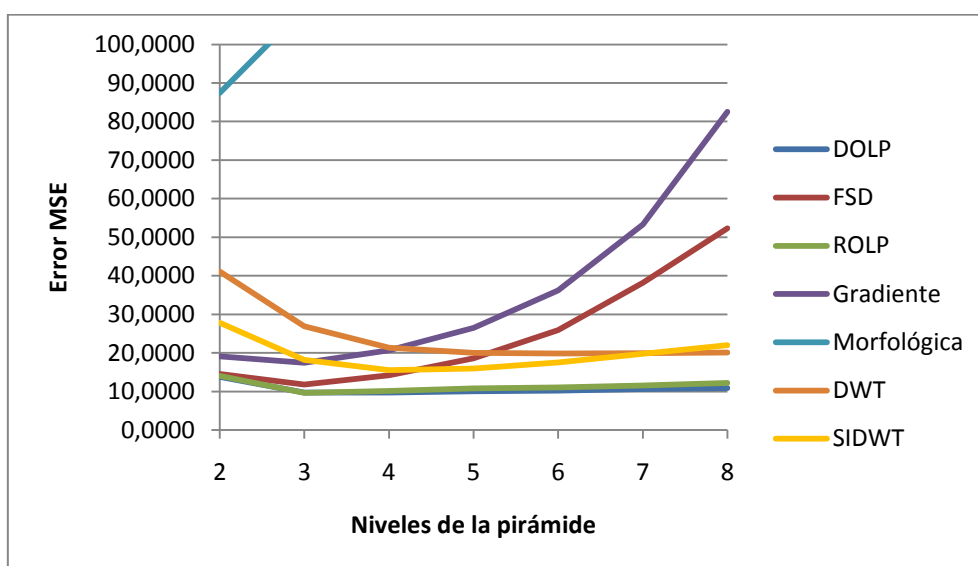


Figura 24: Error MSE en función del número de niveles

A partir de los resultados mostrados en la Figura 24 podemos observar que, en función del método, el número de niveles óptimo resulta ser diferente. Así, para DOLP, ROLP y FSD el número óptimo de niveles es tres. Para SIDWT el menor error se consigue con cuatro niveles, para DWT con seis niveles y para el gradiente con dos. Por último, para la pirámide morfológica, al pasar de dos niveles, el error se dispara considerablemente.

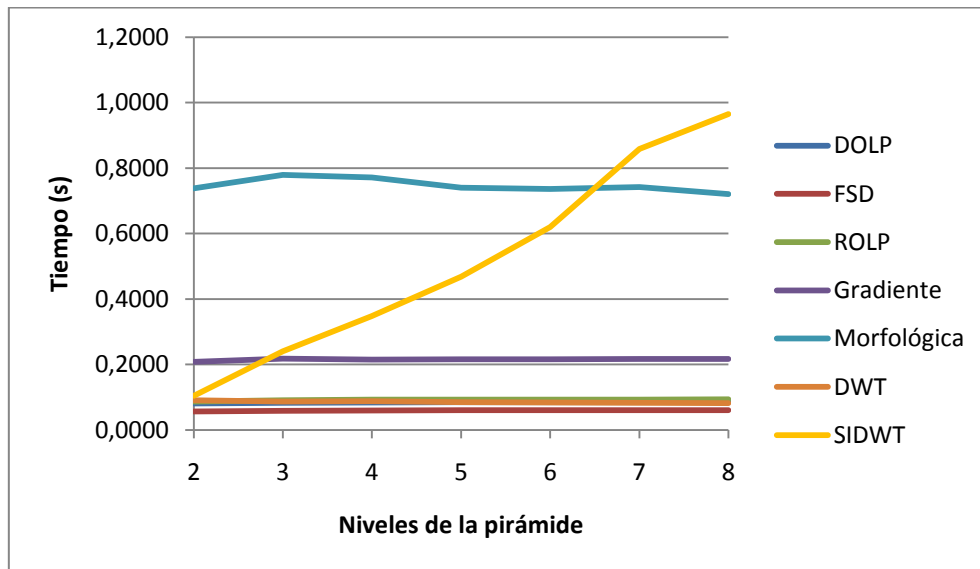


Figura 25: Tiempo en función del número de niveles

Si analizamos el tiempo necesario en función del número de niveles de la pirámide, podemos ver que este no resulta ser un factor relevante en la mayoría de los métodos. Esto se debe a que el mayor consumo de tiempo se produce en el primer nivel de la pirámide, los siguientes niveles son mucho más pequeños y por tanto el tiempo que se emplea en ellos es despreciable. En SIDWT, sin embargo, al ser todos los niveles del mismo tamaño el tiempo aumenta de modo considerable.

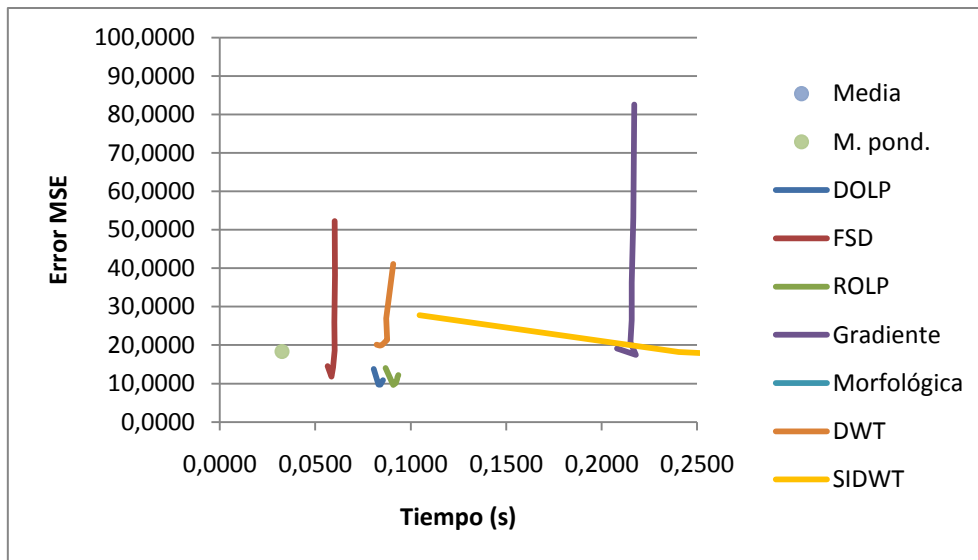


Figura 26: Error cometido en función del tiempo

A partir de los resultados mostrados en la Figura 26 podemos ver que el método que ha ofrecido una mejor calidad ha sido la pirámide ROLP, seguida muy de cerca por la pirámide DOLP. En cuanto al resto de métodos podemos deducir que la pirámide FSD es realmente más rápida que la DOLP (alrededor de un 30%), aunque también es cierto que el error final es mayor, lo que se hace especialmente acusado en pirámides de más de cinco niveles de descomposición, donde el error se duplica. Los métodos más rápidos, como cabía esperar, son la media y la media ponderada, ya que necesitaban muchas menos operaciones que el resto, aunque con ambos se han obtenido errores mayores que con DOLP, ROLP y FSD.

La pirámide del gradiente ha obtenido resultados bastante irregulares, fusionando bien algunas imágenes pero obteniendo errores muy altos en otras, lo que ha generado como promedio un error relativamente alto. La pirámide DWT ha tenido un comportamiento intermedio, con un error ligeramente mayor que el gradiente si bien con un tiempo de cómputo considerablemente menor. La pirámide SIDWT otorga un mejor control sobre el resultado, permitiendo obtener un error alto en tiempo reducido o un error más pequeño empleando más tiempo de cómputo. El peor resultado ha sido obtenido con la pirámide morfológica, tanto en tiempo como en error cometido, ni siquiera entró en la escala de representación del gráfico.

Todos estos resultados se muestran en la Tabla 1 del Anexo II.

5.2. Otros métodos de combinación de coeficientes

Utilizando el mismo conjunto de imágenes que en la sección anterior estudiaremos los efectos que tiene sobre el error final el uso de métodos alternativos de combinación de coeficientes. En este caso, sin embargo, no estudiaremos todos los métodos de fusión, ya que ni en la media simple ni en la media ponderada existe ningún tipo de combinación de coeficientes. En el resto de métodos ahora utilizaremos la media ponderada general de las ecuaciones (48) a (50) con $\alpha = 0,5$ como método de combinación para los coeficientes de baja frecuencia y la media ponderada adaptativa de las ecuaciones (51) y (52) con $\alpha = 3$ para los de alta frecuencia.

En este caso los resultados se recogen en las dos figuras que aparecen seguidamente. La Figura 27 muestra para los nuevos métodos el error obtenido en función del número de niveles de la pirámide, mientras que la Figura 28 muestra el error en función del tiempo.

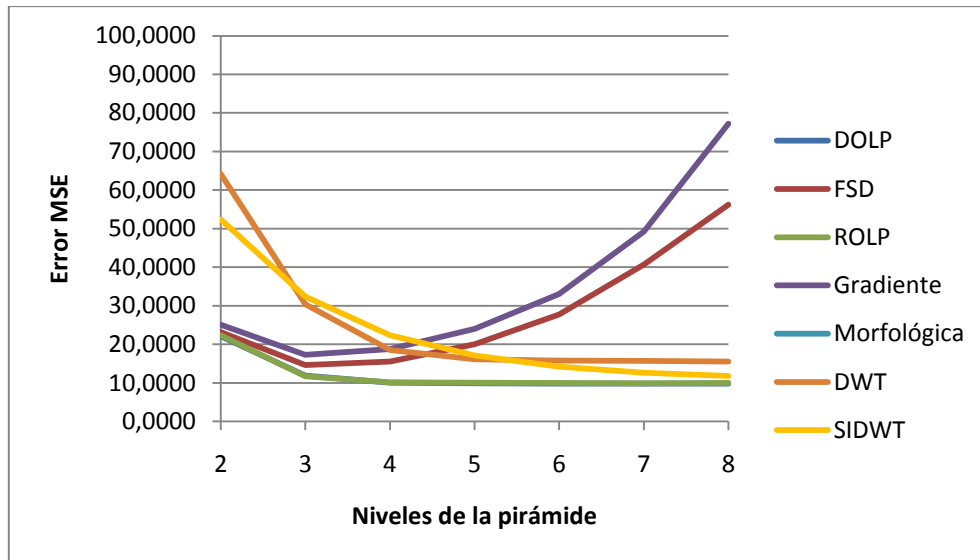


Figura 27: Error en función del número de niveles con combinación de coeficientes basada en la media ponderada

Como puede apreciarse a partir de los resultados mostrados en la Figura 27, con todos los métodos, el error obtenido para la pirámide de dos niveles ha empeorado. Para las pirámides con más niveles, DWT y SIDWT han experimentado una importante mejora en los niveles altos; DOLP, ROLP y la pirámide del gradiente apenas han experimentado cambios, mientras que FSD ha empeorado ligeramente; finalmente la pirámide morfológica ha aumentado considerablemente su error.

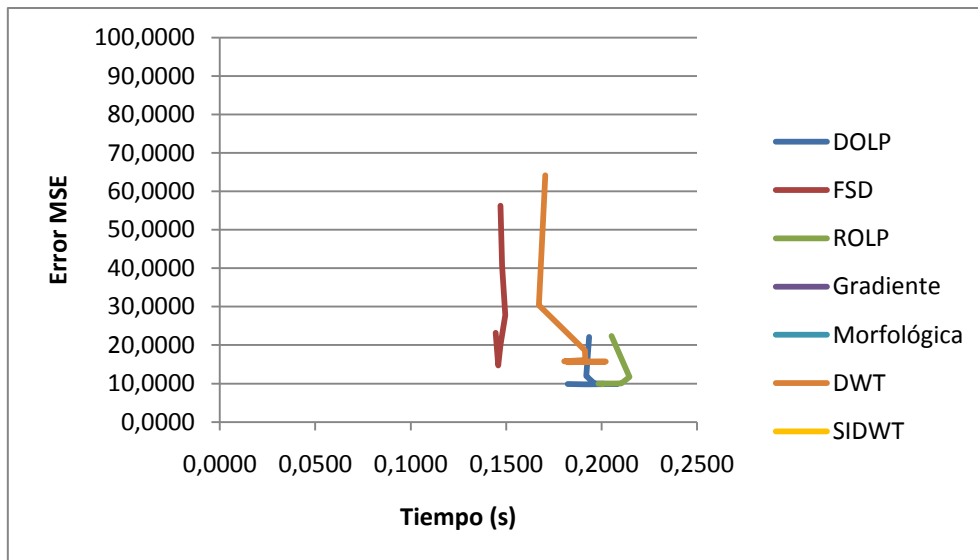


Figura 28: Error en función del tiempo con combinación de coeficientes basada en la media ponderada

A partir de los resultados mostrados en la Figura 28 podemos ver cómo estos nuevos métodos de selección de coeficientes han aumentado bastante el tiempo de cálculo necesario, llegando a obtenerse en algunos casos (pirámide del gradiente, pirámide morfológica y SIDWT) valores que sobrepasan la escala de la gráfica.

Teniendo en cuenta los resultados del error así como el tiempo extra necesario utilizado, el uso de estos métodos alternativos de combinación de coeficientes sólo resulta interesante si se van a usar mediante una transformada de wavelets DWT.

En la Tabla 2 del Anexo II se recogen los resultados de estas pruebas para los diferentes métodos analizados.

Por último a continuación se muestran los resultados de utilizar los métodos de combinación basados en el gradiente. Igual que en el caso anterior haremos las pruebas sobre las pirámides DOLP, FSD, ROLP, del gradiente, morfológica, DWT y SIDWT.

La Figura 29 muestra el error obtenido para cada nivel de la pirámide, mientras que la Figura 30 muestra la relación entre el tiempo necesario y el error final.

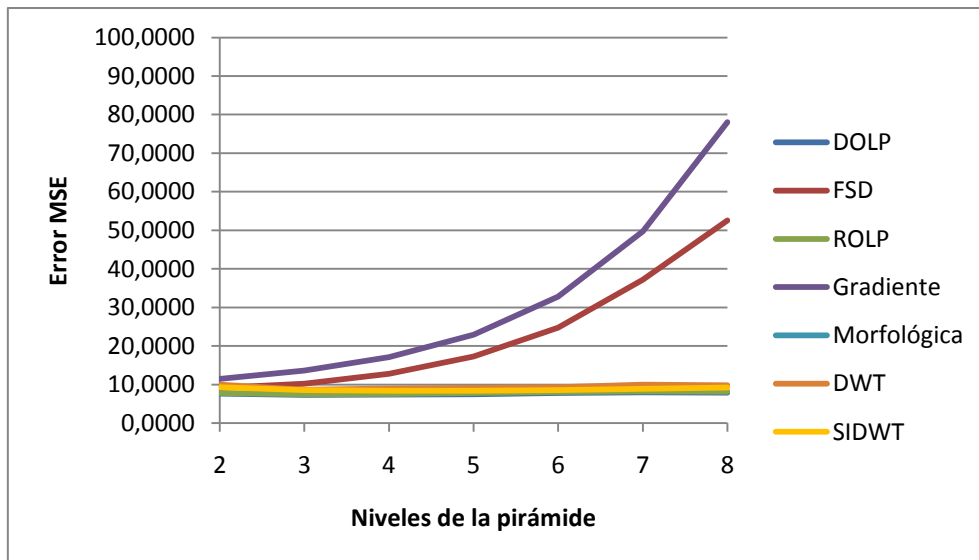


Figura 29: Error en función del número de niveles con combinación de coeficientes basada en el gradiente

En el gráfico de la Figura 29 podemos ver que gracias a los nuevos coeficientes que han intervenido en la fusión obtenidos a partir del gradiente, se han conseguido mejores resultados para casi todos los métodos, con la única excepción de la pirámide morfológica. Llama especialmente la atención el hecho de que el error con una pirámide de únicamente dos niveles de descomposición ha bajado hasta situarse alrededor de 10, descendiendo en pirámides de más niveles para los métodos DOLP, ROLP, DWT y SIDWT y subiendo en FSD y la pirámide morfológica.

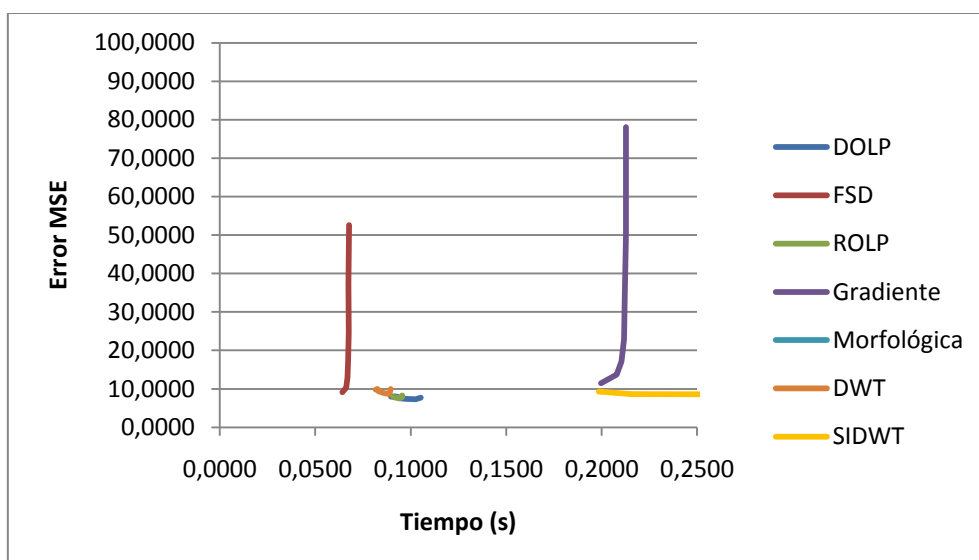


Figura 30: Error en función del tiempo con combinación de coeficientes basada en el gradiente

A la vista del error obtenido en función del tiempo, mostrado en la gráfica de la Figura 30, observamos que todos los métodos han aumentado su tiempo en mayor medida, aunque en todos los casos mucho menos que con las medias ponderadas que vimos anteriormente. Por tanto, podemos concluir que los métodos de combinación de coeficientes basados en el gradiente son una buena elección.

En la Tabla 3 del Anexo II se muestran los resultados de este nuevo método de selección de coeficientes.

En esta sección hemos podido comprobar que el método de selección de coeficientes que ha dado mejores resultados, a costa de aumentar ligeramente el tiempo de procesamiento, ha sido el basado en el gradiente local.

Otra conclusión que podemos extraer del análisis previo es que ningún método de selección de coeficientes de los estudiados aquí ha dado buenos resultados con la pirámide morfológica, lo cual puede deberse a que ésta es la única pirámide en la que no entran en juego los conceptos de “altas” y “bajas” frecuencias. En cualquier caso su elevado tiempo de procesamiento parece sugerir que no es un método adecuado para su uso en determinados sistemas, particularmente en sistemas empotrados.

5.3. Fusión en imágenes de color

Como vimos en el Capítulo 3 vamos a estudiar dos métodos diferentes para este tipo de fusión. En primer lugar analizaremos el efecto de realizar la fusión sobre los canales R, G y B por separado para a continuación estudiar el error que se produce al convertir la imagen a YIQ y realizar la fusión sobre el canal de luminancia, mezclando los otros canales con (56). En todos los casos el método de selección de coeficientes es el del máximo para las altas frecuencias y la media para las bajas.

Para la primera prueba nuevamente estudiaremos todos los métodos que usamos en el Capítulo 2, aunque en esta ocasión las imágenes sobre las que realizaremos las pruebas serán otras, ya que las anteriores estaban en escala de grises. La Figura 31 muestra tres imágenes de muestra representativas del nuevo conjunto constituido por veinte imágenes sobre el que se ha realizado el análisis.



Figura 31: Algunas de las imágenes en color que se van a usar en las pruebas

Tras aplicar los métodos sobre todas las imágenes obtenemos los resultados de las siguientes figuras. En la Figura 32 podemos observar el error obtenido para cada uno de los métodos en función del tiempo necesario.

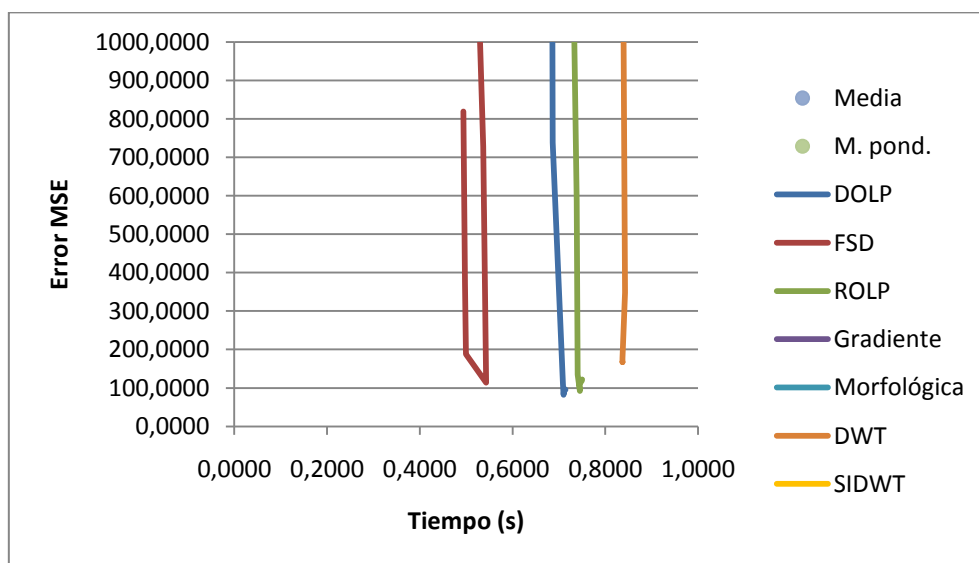


Figura 32: Error cometido en función del tiempo para fusión RGB

En la Figura 32 se puede observar que en la fusión de imágenes canal a canal, el método más rápido sigue siendo FSD, seguido de DOLP, ROLP y DWT. El método con un menor error ha sido la pirámide DOLP de 5 niveles, seguida de ROLP, FSD y DWT. El resto de métodos han tenido un error o un tiempo de procesamiento sensiblemente mayor a éstos, por lo que no se han tenido en cuenta. La Tabla 4 del Anexo II recoge estos datos de forma numérica.

Ahora, utilizando el mismo conjunto de imágenes, realizaremos las pruebas utilizando la conversión a YIQ. La Figura 33 muestra el error frente al tiempo en el nuevo escenario.

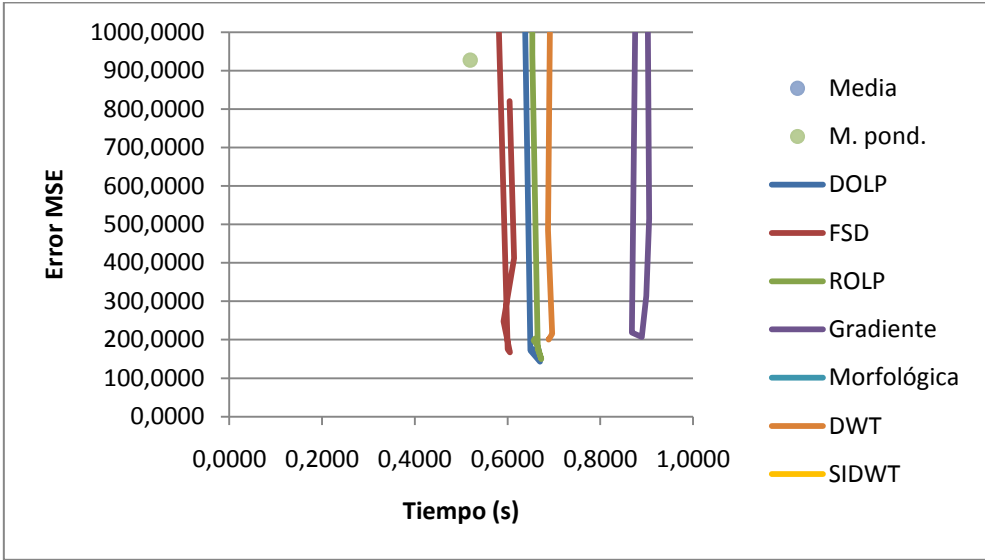


Figura 33: Error cometido en función del tiempo para fusión YIQ

Los nuevos resultados muestran una reducción en el tiempo de procesamiento necesario para todos los métodos excepto para FSD. El método con un menor error también ha sido DOLP en este caso, aunque ha doblado el error obtenido con RGB por lo que no se puede considerar una buena solución. Todos los datos se encuentran en la Tabla 5 del Anexo II.

Tras comprobar que la fusión sobre el canal Y en YIQ ha resultado más rápida que la fusión canal a canal de RGB, esto lleva a probar si el tiempo ganado al realizar la fusión en un solo canal se puede aprovechar para utilizar una combinación de coeficientes más compleja, como el método basado en el gradiente global presentado en la sección 2.11. La Figura 34 muestra el error frente al tiempo al usar este método.

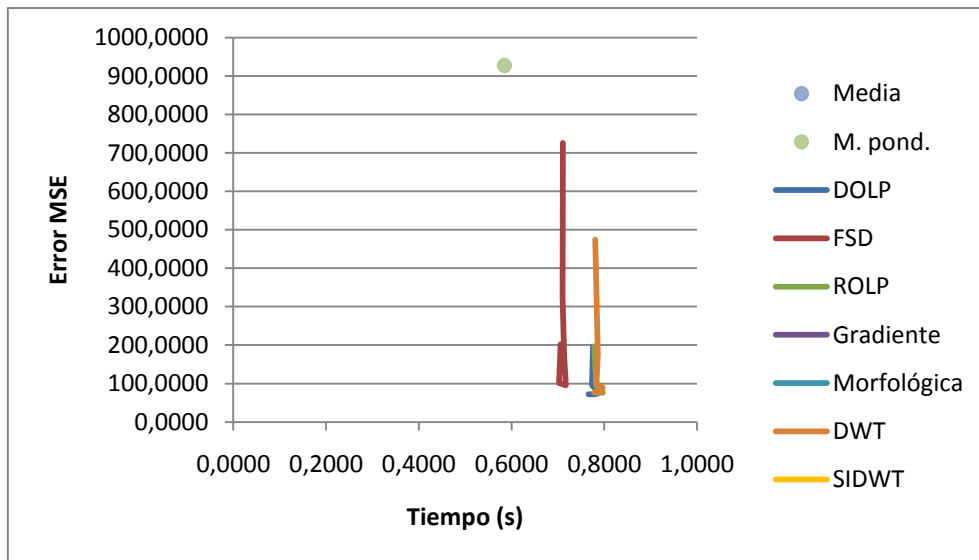


Figura 34: Error cometido en función del tiempo para YIQ y combinación basada en gradiente

En la Figura 34 se puede observar que el error es ahora menor que el que se obtuvo en la fusión RGB canal a canal. Sin embargo todos los métodos han aumentado mucho su tiempo de procesamiento, algunos incluso por encima de la fusión canal a canal, por lo que no se puede afirmar que esta estrategia sea claramente interesante. Los datos en detalle se encuentran en la Tabla 6 del Anexo II.

Por último se comprobará si al utilizar en la fusión canal a canal el método de combinación de coeficientes basado en el gradiente el error también disminuye, o por el contrario el valor que se ha conseguido con YIQ es el mínimo que podemos encontrar.

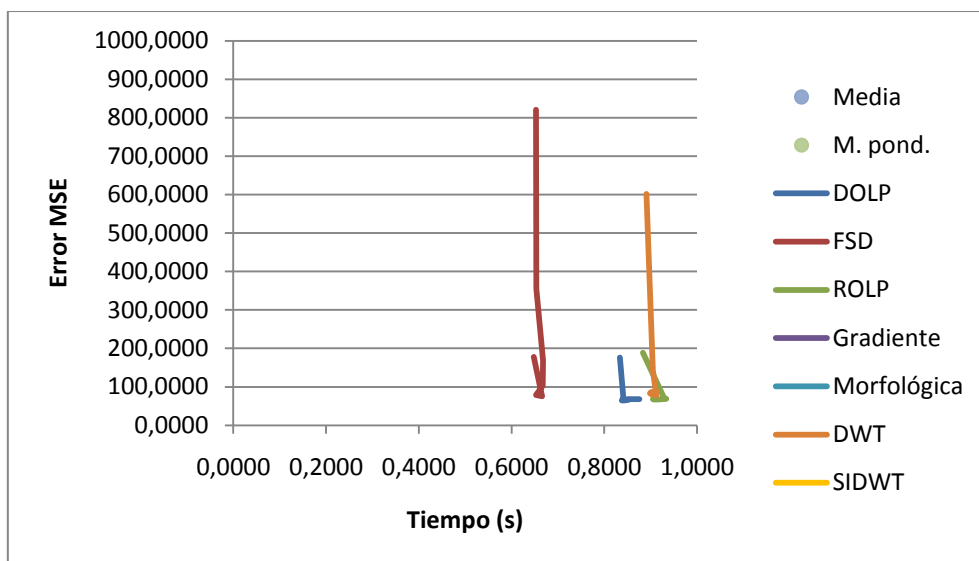


Figura 35: Error cometido en función del tiempo para RGB y combinación basada en gradiente

La Figura 35 muestra que la pirámide DOLP ha obtenido un error alrededor de un 11% menor al de la Figura 34 a costa de aumentar un 7% el tiempo necesario. Los métodos FSD, ROLP y DWT también han reducido su error y todos salvo FSD han aumentado su tiempo. La Tabla 7 del Anexo II recoge todos estos datos.

A la vista de todos los resultados en el caso de fusión en imágenes de color, el tiempo está relacionado directamente con la calidad que se obtendrá. El método más rápido, YIQ, también fue el que obtuvo unos peores resultados, mientras que RGB con combinación de coeficientes basada en el máximo obtuvo un error sensiblemente menor.

5.4. Fusión en imágenes con dimensiones inadecuadas

Para realizar las pruebas de esta sección volveremos a usar las imágenes de la sección 5.1, aunque como estas imágenes tenían unas dimensiones de 512×512 píxeles que permitían su fusión sin problemas, tendremos que recortarlas para poder aplicar alguna de las técnicas vistas en el Capítulo 4. En concreto nos quedaremos con la región de 497×497 píxeles de la parte superior izquierda de las imágenes.

En primer lugar veremos la respuesta de los métodos de fusión al eliminar el paso de redimensionado de los métodos. En este caso el método DOLP se definirá como vimos en las ecuaciones (58) a (60), el método ROLP se definirá de modo análogo pero cambiando las sumas por productos y las rectas por divisiones, y el método del gradiente utilizará la nueva pirámide gaussiana. La pirámide morfológica será igual que en las ecuaciones (28) a (30), salvo que omitiendo la eliminación y reconstrucción de filas y columnas. El método DWT se definirá igual que el SIDWT pero ignorando el cambio de la máscara en función del nivel. El método SIDWT se incluye a modo de referencia. En este caso no se podrá estudiar el método FSD ya que al eliminar la reducción de tamaño quedaría igual que el método DOLP.

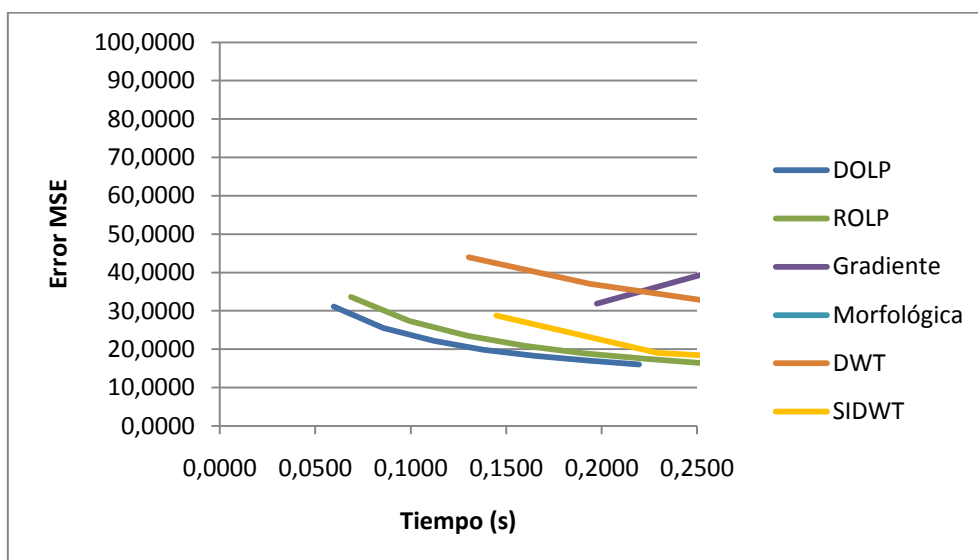


Figura 36: Error cometido en función del tiempo sin re-escalado

La Figura 36 muestra que usando esta técnica en casi todos los métodos existe una fuerte correspondencia entre el tiempo y el error. El método con un mejor comportamiento ha sido DOLP, y llama la atención que pese a lo parecidos que son los métodos DWT y SIDWT este último obtenga un error sensiblemente inferior. Los datos en detalle se encuentran en la Tabla 8 del Anexo II.

A continuación veremos los efectos de añadir bandas negras a las imágenes hasta alcanzar el tamaño necesario para poder aplicar los métodos originales. En función del número de niveles que vaya a tener la pirámide el tamaño variará: así, para una pirámide de 2 niveles deberá ser 500×500 , para una de 3 será 504×504 y para una de entre 4 y 8 niveles el tamaño deberá ser 512×512 . Aunque el método SIDWT podría realizar la fusión sin problemas, para este método también se añadirán las bandas negras como referencia.

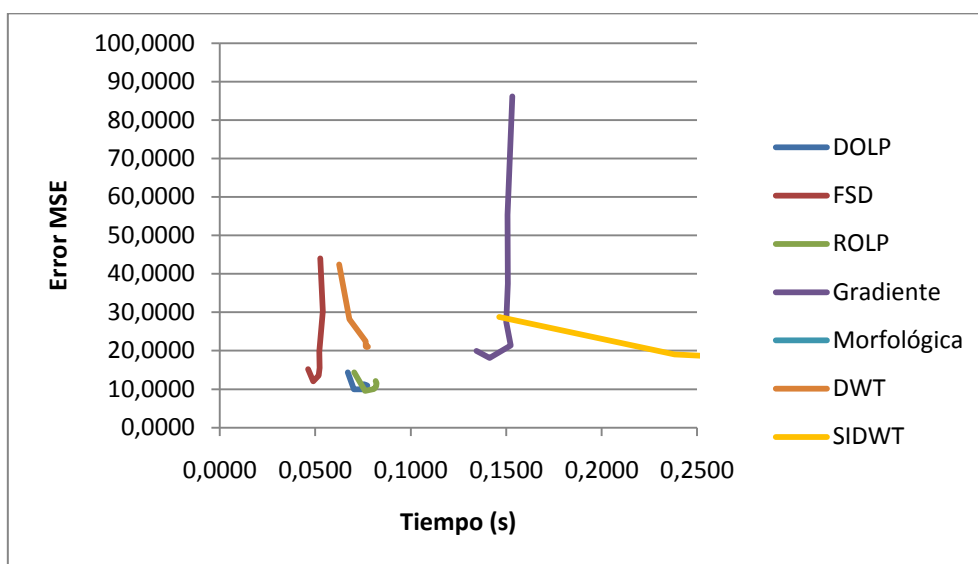


Figura 37: Error cometido en función del tiempo añadiendo bandas negras

En la Figura 37 podemos observar que por lo general tanto el error cometido como el tiempo necesario han sido mucho menores que en el caso anterior. Como excepción está el método SIDWT, que no ha experimentado cambios. En este caso el método con menor error ha sido ROLP, aunque la diferencia con DOLP es prácticamente nula, mientras que el método más rápido ha sido FSD. Todos los datos se encuentran en la Tabla 9 del Anexo II.

Por último se probará a redimensionar las imágenes usando algoritmos de re-escalado. Las dimensiones finales serán, en todos los casos, las mismas que se mencionaron al hablar de las bandas negras. Conviene destacar que este redimensionado sólo se realiza una vez, antes de proceder a la fusión. Esto es, si se va a realizar una fusión de 3 niveles la imagen original se re-escalará a 504×504 píxeles directamente y no a 500×500 para la fusión del primer nivel y 504×504 para el siguiente. Al igual que en el caso anterior, en el método SIDWT haremos este re-escalado a pesar de no ser necesario.

El primer método de re-escalado es interpolación por vecino más próximo, con el que obtenemos los siguientes resultados.

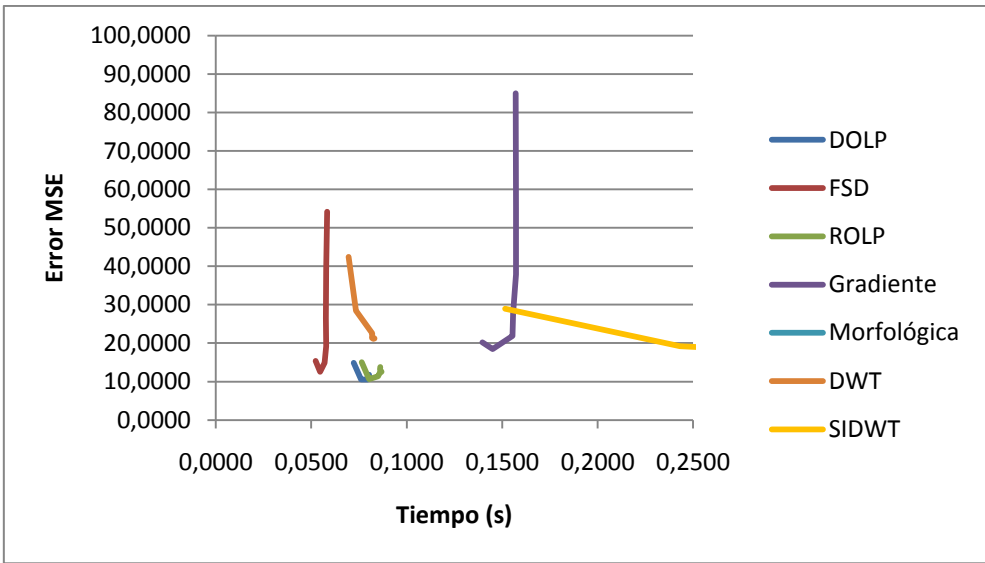


Figura 38: Error cometido en función del tiempo con re-escalado por vecino más próximo

La Figura 38 muestra que usando la interpolación por vecino más próximo los resultados son muy parecidos a los obtenidos al añadir barras. Al observar en detalle se descubre que en realidad en este caso el tiempo ha sido ligeramente mayor y que el error también ha aumentado. En la Tabla 10 del Anexo II se pueden revisar los datos correspondientes a esta prueba.

El segundo y último método será la interpolación bicúbica. Con este método obtenemos los siguientes resultados.

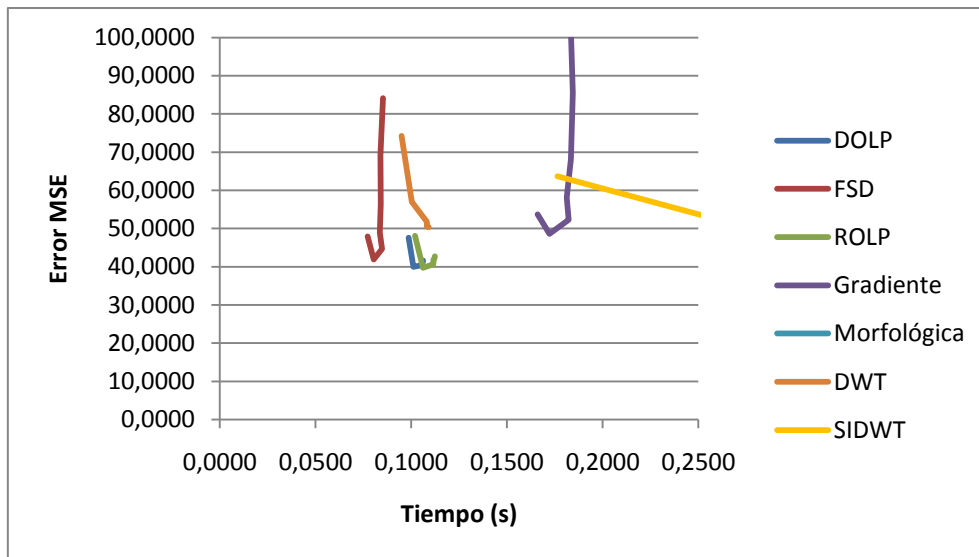


Figura 39: error cometido en función del tiempo con re-escalado bicúbico

En la Figura 39 se observa que el tiempo necesario ha aumentado considerablemente respecto a la interpolación por vecino más próximo. Igualmente se puede advertir que este tiempo extra no ha reducido el error, que es ahora mucho mayor. Estos mismos datos se muestran en la Tabla 11 del Anexo II.

En este caso la conclusión es clara: el mejor procedimiento, tanto por error como por tiempo, consiste en añadir bandas negras a las imágenes hasta completar las dimensiones necesarias. El resto de los métodos estudiados no aporta ninguna ventaja, por lo que los resultados obtenidos con ellos no es necesario mostrarlos.

Capítulo 6

Conclusiones y trabajo futuro

Hemos realizado una revisión de métodos para fusión de imágenes, habiendo extendido los mecanismos de fusión para el caso de imágenes en color. Se ha analizado el comportamiento de cada una de las estrategias estudiadas.

Las principales conclusiones a las que hemos llegado son:

- 1) El método que ha obtenido un menor error respecto a la imagen de referencia ha sido la pirámide DOLP de cuatro niveles, usando combinación de coeficientes basada en el gradiente global.
- 2) En el caso de las imágenes en color el mejor método ha resultado ser la fusión RGB canal a canal.
- 3) Para las imágenes con dimensiones inadecuadas el menor error se ha conseguido añadiendo bandas negras a la imagen hasta completar el tamaño necesario. Esto es debido principalmente a la naturaleza simétrica de las operaciones de filtrado.
- 4) El método computacionalmente más eficiente en términos de tiempos de ejecución es el basado en pirámide FSD de dos niveles con combinación simple de coeficientes.

En las imágenes mostradas en la Figura 40, se muestran los resultados de aplicar respectivamente los métodos DOLP y FSD a las imágenes de color mostradas en la Figura 18.

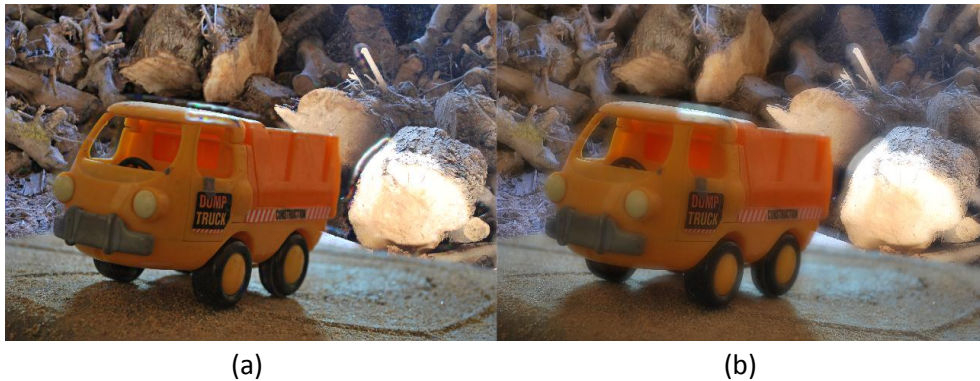


Figura 40: (a) Resultados de una fusión usando los métodos (a) DOLP; (b) FSD

En la Figura 40 se observa cómo, en efecto, el método DOLP que hemos identificado como el de menor error obtiene un resultado mucho más nítido que el método más rápido FSD. Sin embargo, sobre ambas imágenes todavía resultan apreciables algunos artefactos como los comentados en la sección 2.3.

A partir de las consideraciones anteriores conviene resaltar las líneas de investigación identificadas, a partir de los experimentos realizados, con proyección de futuro:

- 1) Es necesario profundizar en la búsqueda de métodos de fusión que eliminen o al menos reduzcan en la medida de lo posible los artefactos que aparecen en mayor o menor medida con los métodos de fusión.
- 2) Convendría analizar las posibilidades de implementación de los métodos en dispositivos hardware con vistas a su posible aplicación en los sistemas de captura de imágenes en el futuro para llegar a tiempos de proceso identificables con tiempos reales.
- 3) Como se ha comentado en la introducción para incluir estas técnicas en una cámara real habría que investigar sobre las posibilidades de aplicación de métodos automáticos de enfoque digital que permitan conocer las tomas que se deben capturar en cada caso, así como técnicas de registro de imágenes para alinear correctamente las imágenes en caso de que la cámara se moviese durante la realización de las mismas.

Finalmente, abriendo una línea de investigación futura, conviene comentar un resultado interesante que se ha observado de modo accidental durante la realización de las pruebas es que las fusiones en las que las distintas tomas tenían diferente tiempo de exposición producían resultados bastante satisfactorios. Tal es el caso mostrado en la Figura 41 donde se muestran dos imágenes con distinto enfoque y distinto tiempo de exposición, en (a) aparece la calle y en (b) el castillo.

Tras fundir ambas imágenes se obtiene el resultado mostrado en la Figura 42. En esta nueva imagen, se puede observar que se han conservado los detalles tanto de las zonas oscuras como de las zonas claras. La imposibilidad de realizar un análisis cualitativo de este tipo de fusión hace difícil presentar unos resultados concluyentes, sin embargo éste resulta un campo interesante para continuar la investigación ya que al menos cualitativamente los resultados que aparecen resultan prometedores.

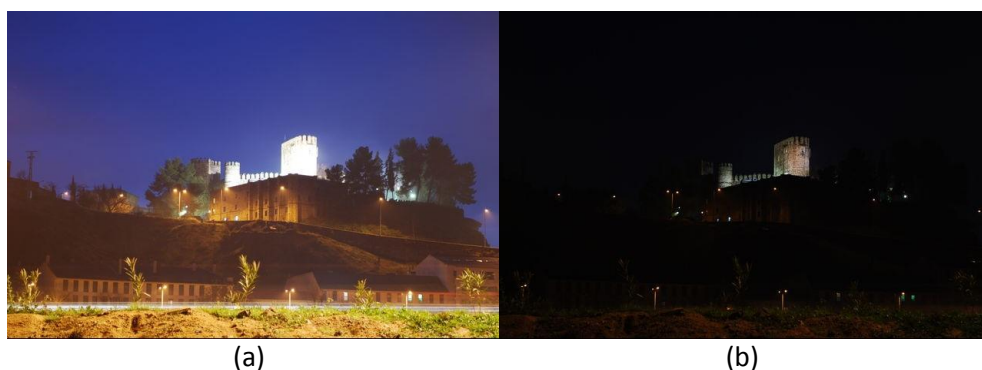


Figura 41: Dos imágenes con distintos tiempos de exposición.



Figura 42: Fusión de las dos imágenes con distinta exposición mostradas en la Figura 41

Anexo I

Altas y bajas frecuencias en imágenes

En teoría de señales es habitual el uso de los conceptos de altas y bajas frecuencias. La velocidad con la que una determinada señal cambia se mide en ciclos o Hertzios y recibe el nombre de “frecuencia de la señal”. La Figura 43 muestra dos señales sinusoidales diferentes, la de la izquierda con una frecuencia de 10 Hertzios y la de la derecha de 100 Hz.

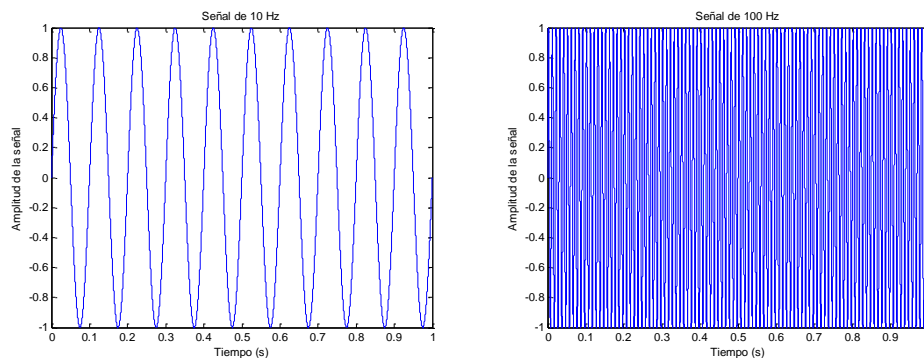


Figura 43: Dos señales de distinta frecuencia

Una señal será de alta frecuencia si varía con rapidez, mientras que si cambia más lentamente será de baja frecuencia. Los límites que determinan de qué tipo es una señal dependerán de su tipo, de este modo tenemos que 40 kHz es una frecuencia muy alta cuando hablamos de un sonido pero muy baja si hablamos de ondas de radio.

En las imágenes anteriores la variación de la señal se produce en función del tiempo. En el caso de las imágenes nos encontramos ante señales bidimensionales donde las variaciones en lugar de ser temporales son espaciales. Por tanto, para identificar las zonas de altas y bajas frecuencias hay que referirse al dominio espacial en lugar del temporal. Dicho dominio espacial es doble, esto es, se produce tanto en el eje x como en el y. Sin embargo, para clarificar los conceptos se puede utilizar una aproximación que facilita esta identificación si suponemos que tratamos con imágenes en una única dimensión, esto es unidimensionales. Tal es el caso representado en la imagen de la Figura 44 donde se ha trazado una línea en la dirección x, la cual permite determinar lo que se conoce como perfil de intensidad (Pajares & de la Cruz, 2001).



Figura 44: Dos imágenes con su línea central remarcada

Pues bien, si tomamos como referencia la línea marcada, podemos considerar el perfil que genera como una señal unidimensional que toma valores entre 0 (correspondiendo a intensidades cuya componente es negra) y 255 (con componente blanca). La Figura 45 muestra la forma de esta señal para ambas imágenes en la Figura 44.

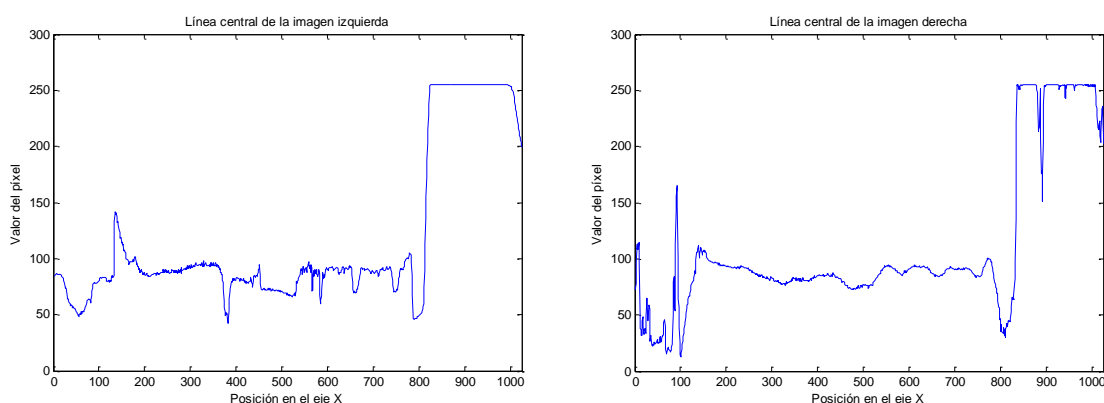


Figura 45: Perfil de las imágenes

En la Figura 45 puede observarse cómo ambas señales presentan aproximadamente tres zonas bien diferenciadas, recorriendo todo el rango de la imagen en la dirección x: a) desde el nivel de intensidad 0 hasta el 100; b) desde el 100 hasta el 800 y c) desde el 800 hasta el 1024. Las zonas a) y c) presentan una mayor variabilidad en el perfil de la derecha que en el de la izquierda, mientras que la zona b) varía de una forma más brusca en el de la izquierda. Esta división o variabilidad es la misma que se manifiesta al estudiar el enfoque de las imágenes, lo que indica claramente que las zonas enfocadas se corresponden con las frecuencias mayores y viceversa, las zonas con perfiles suaves, esto es con escasa variabilidad, se corresponden con las zonas desenfocadas.

Por último en la Figura 46 se muestra la imagen resultante de fusionar las imágenes de la Figura 44 junto con el perfil correspondiente a su línea central.

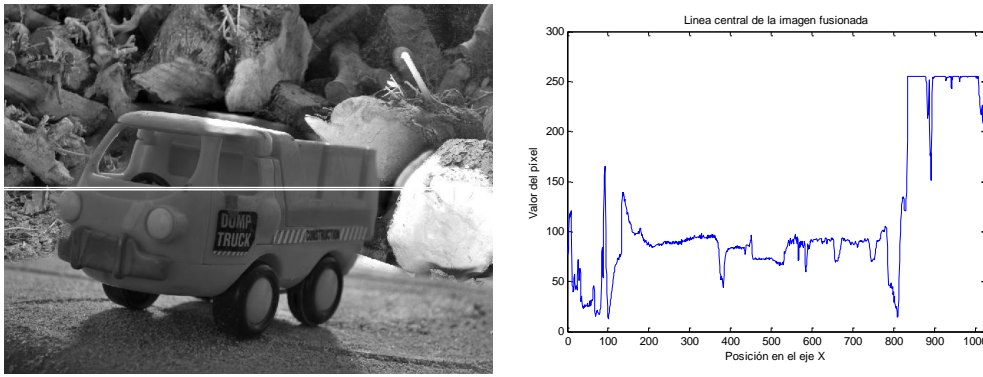


Figura 46: Imagen fusionada y el perfil de su línea central

En el perfil de intensidad de la imagen fusionada se aprecia que la forma de la señal imita en cada zona a la imagen original asociada con las mayores frecuencias. Esto se puede observar con mayor facilidad en la Figura 47.

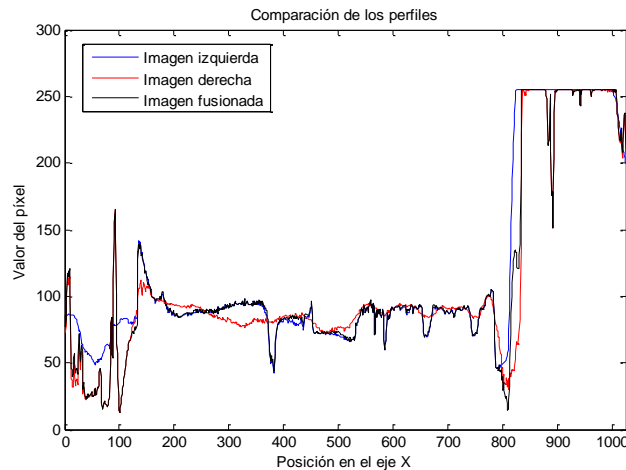


Figura 47: Comparación de los perfiles de las imágenes

En efecto, en la mencionada Figura 47 se puede observar cómo la señal correspondiente a la imagen fusionada copia en cada momento el perfil de la señal con mayor frecuencia. En las zonas de transición, sin embargo, toma un valor intermedio.

Para extrapolar el concepto de frecuencia visto aquí a imágenes bidimensionales, se utilizan diferentes aproximaciones. Así, la pirámide Laplaciana descrita en la sección 2.4, junto a los métodos que la utilizan (pirámide FSD, pirámide ROLP), se basan en desenfocar la imagen para quedarse con las bajas frecuencias y posteriormente eliminar de la imagen original estas bajas frecuencias con el fin de hallar las altas. El método del gradiente descrito en la sección 2.7 comprueba cómo varía la imagen a lo largo de cuatro direcciones distintas para ver las frecuencias presentes en cada una de ellas. Por último la pirámide DWT de la sección 2.9, junto con su variación SIDWT, se basa en observar la variación de la imagen a lo largo de los ejes horizontal y vertical.

Anexo II

Tablas de resultados

Tabla 1: Resultados de la fusión de imágenes con selección simple de coeficientes.

Método	MSE	Tiempo
Media	114,0530	0,0058
M. ponderada	18,3337	0,0326
DOLP(2)	13,8306	0,0807
DOLP(3)	9,7020	0,0831
DOLP(4)	9,7231	0,0842
DOLP(5)	10,0764	0,0845
DOLP(6)	10,2671	0,0848
DOLP(7)	10,6375	0,0851
DOLP(8)	10,9021	0,0855
FSD(2)	14,5307	0,0565
FSD(3)	11,8185	0,0585
FSD(4)	14,1943	0,0593
FSD(5)	18,6409	0,0602
FSD(6)	25,9436	0,0601
FSD(7)	38,2098	0,0603
FSD(8)	52,2897	0,0601
ROLP(2)	14,0746	0,0869
ROLP(3)	9,6299	0,0909
ROLP(4)	10,1517	0,0922
ROLP(5)	10,8431	0,0927
ROLP(6)	11,1108	0,0928
ROLP(7)	11,6084	0,0931
ROLP(8)	12,2055	0,0935
Gradiente(2)	19,1159	0,2081
Gradiente(3)	17,4431	0,2180
Gradiente(4)	20,6732	0,2151

Método	MSE	Tiempo
Gradiente(5)	26,5293	0,2158
Gradiente(6)	36,2317	0,2157
Gradiente(7)	53,2903	0,2167
Gradiente(8)	82,5420	0,2171
Morfológica(2)	87,3950	0,7381
Morfológica(3)	109,4240	0,7788
Morfológica(4)	145,2417	0,7711
Morfológica(5)	190,1333	0,7397
Morfológica(6)	279,7267	0,7360
Morfológica(7)	289,6767	0,7424
Morfológica(8)	457,4433	0,7203
DWT(2)	41,1397	0,0909
DWT(3)	26,9427	0,0871
DWT(4)	21,3960	0,0876
DWT(5)	20,0247	0,0849
DWT(6)	19,8617	0,0840
DWT(7)	19,9387	0,0832
DWT(8)	20,0787	0,0821
SIDWT(2)	27,8143	0,1047
SIDWT(3)	18,1969	0,2404
SIDWT(4)	15,5920	0,3481
SIDWT(5)	15,9470	0,4683
SIDWT(6)	17,5810	0,6199
SIDWT(7)	19,7943	0,8581
SIDWT(8)	22,0463	0,9648

Tabla 2: Resultados de la fusión usando métodos de combinación basados en la media ponderada.

Método	MSE	Tiempo
DOLP(2)	22,1285	0,1935
DOLP(3)	11,9215	0,1919
DOLP(4)	10,0678	0,1962
DOLP(5)	9,8797	0,2084
DOLP(6)	9,8200	0,1932
DOLP(7)	9,8084	0,1914
DOLP(8)	9,8462	0,1822
FSD(2)	23,2207	0,1445
FSD(3)	14,6756	0,1459
FSD(4)	15,5469	0,1461
FSD(5)	20,0726	0,1471
FSD(6)	27,7386	0,1495
FSD(7)	40,7123	0,1478
FSD(8)	56,2223	0,1470
ROLP(2)	22,3790	0,2052
ROLP(3)	11,6995	0,2146
ROLP(4)	10,1360	0,2106
ROLP(5)	10,0396	0,2104
ROLP(6)	9,9938	0,2004
ROLP(7)	9,9329	0,2005
ROLP(8)	9,9609	0,1983
Gradiente(2)	25,0827	0,5568
Gradiente(3)	17,3201	0,5600
Gradiente(4)	18,8211	0,5718
Gradiente(5)	24,0249	0,5769

Método	MSE	Tiempo
Gradiente(6)	33,0516	0,5788
Gradiente(7)	49,2891	0,5891
Gradiente(8)	77,1993	0,5784
Morfológica(2)	341,2533	0,7106
Morfológica(3)	370,6967	0,6777
Morfológica(4)	364,6733	0,6625
Morfológica(5)	373,3967	0,7803
Morfológica(6)	475,3600	0,7436
Morfológica(7)	365,1467	0,7553
Morfológica(8)	374,6967	0,7142
DWT(2)	64,1880	0,1705
DWT(3)	30,3997	0,1671
DWT(4)	18,5810	0,1913
DWT(5)	16,1337	0,1914
DWT(6)	15,7997	0,1804
DWT(7)	15,7093	0,2022
DWT(8)	15,5527	0,1823
SIDWT(2)	52,3830	0,8447
SIDWT(3)	32,3807	1,1619
SIDWT(4)	22,3934	1,2858
SIDWT(5)	17,1420	1,6451
SIDWT(6)	14,2412	2,0254
SIDWT(7)	12,6600	2,4498
SIDWT(8)	11,7806	2,6325

Tabla 3: Resultados de la fusión usando métodos de combinación basados en el gradiente.

Método	MSE	Tiempo
DOLP(2)	7,6836	0,1053
DOLP(3)	7,3383	0,1026
DOLP(4)	7,4015	0,0981
DOLP(5)	7,5377	0,0943
DOLP(6)	7,8539	0,0936
DOLP(7)	8,0300	0,0916
DOLP(8)	7,9247	0,0896
FSD(2)	9,1713	0,0643
FSD(3)	10,2568	0,0662
FSD(4)	12,8376	0,0668
FSD(5)	17,2857	0,0673
FSD(6)	24,7348	0,0675
FSD(7)	37,2389	0,0674
FSD(8)	52,5680	0,0677
ROLP(2)	7,8626	0,0909
ROLP(3)	7,5394	0,0935
ROLP(4)	7,5794	0,0942
ROLP(5)	7,8552	0,0955
ROLP(6)	8,1232	0,0953
ROLP(7)	8,2668	0,0955
ROLP(8)	8,1880	0,0956
Gradiente(2)	11,4712	0,1997
Gradiente(3)	13,6729	0,2079
Gradiente(4)	17,1587	0,2105
Gradiente(5)	22,9072	0,2117

Método	MSE	Tiempo
Gradiente(6)	32,8156	0,2121
Gradiente(7)	49,7706	0,2128
Gradiente(8)	78,0490	0,2127
Morfológica(2)	270,3833	0,6470
Morfológica(3)	312,8333	0,9384
Morfológica(4)	328,9400	0,9845
Morfológica(5)	375,1867	0,9761
Morfológica(6)	477,4067	1,0561
Morfológica(7)	533,8533	1,0469
Morfológica(8)	899,2833	1,1037
DWT(2)	9,9738	0,0894
DWT(3)	8,6705	0,0888
DWT(4)	8,9719	0,0855
DWT(5)	9,1870	0,0842
DWT(6)	9,3135	0,0834
DWT(7)	9,9765	0,0827
DWT(8)	9,8100	0,0817
SIDWT(2)	9,2951	0,1986
SIDWT(3)	8,6123	0,2158
SIDWT(4)	8,3950	0,4330
SIDWT(5)	8,4441	0,6196
SIDWT(6)	8,6174	0,7341
SIDWT(7)	8,8867	0,9151
SIDWT(8)	9,2578	1,1135

Tabla 4: Resultados de la fusión RGB canal a canal.

Método	MSE	Tiempo
Media	9365,4733	0,0516
M. ponderada	1356,8343	0,2530
DOLP(2)	2907,0633	0,6815
DOLP(3)	738,1910	0,6863
DOLP(4)	110,8677	0,7090
DOLP(5)	81,8557	0,7098
DOLP(6)	91,6897	0,7109
DOLP(7)	95,5803	0,7141
DOLP(8)	96,9347	0,7131
FSD(2)	2915,1867	0,4816
FSD(3)	727,1457	0,5369
FSD(4)	117,6333	0,5427
FSD(5)	113,4177	0,5429
FSD(6)	187,8800	0,4992
FSD(7)	370,7833	0,4971
FSD(8)	818,6400	0,4943
ROLP(2)	2385,2087	0,7161
ROLP(3)	586,3833	0,7380
ROLP(4)	134,8120	0,7406
ROLP(5)	91,7677	0,7451
ROLP(6)	110,4353	0,7447
ROLP(7)	118,6663	0,7491
ROLP(8)	122,1340	0,7498
Gradiente(2)	3064,6157	1,3231
Gradiente(3)	872,4913	1,3976
Gradiente(4)	182,3157	1,4230

Método	MSE	Tiempo
Gradiente(5)	168,8600	1,4027
Gradiente(6)	256,6400	1,3819
Gradiente(7)	454,1733	1,4056
Gradiente(8)	954,0967	1,4064
Morfológica(2)	2842,1733	4,4381
Morfológica(3)	1584,8427	4,5742
Morfológica(4)	1389,1740	4,6900
Morfológica(5)	2263,5933	4,7490
Morfológica(6)	3352,6533	4,7871
Morfológica(7)	3617,4867	4,8936
Morfológica(8)	5654,5867	4,8973
DWT(2)	6800,6267	0,8244
DWT(3)	4098,9347	0,8288
DWT(4)	1646,6993	0,8361
DWT(5)	348,0460	0,8425
DWT(6)	177,3493	0,8373
DWT(7)	166,7613	0,8372
DWT(8)	168,2087	0,8372
SIDWT(2)	5646,2367	1,3876
SIDWT(3)	3620,0497	2,2165
SIDWT(4)	2191,9377	3,1795
SIDWT(5)	1251,6120	4,2974
SIDWT(6)	674,6257	5,4861
SIDWT(7)	359,0617	6,8313
SIDWT(8)	205,9787	8,4286

Tabla 5: Resultados de la fusión YIQ en el canal Y.

Método	MSE	Tiempo
Media	11386,6467	0,4558
M. ponderada	927,0290	0,5195
DOLP(2)	3902,4567	0,6370
DOLP(3)	1040,6933	0,6373
DOLP(4)	171,6740	0,6490
DOLP(5)	142,5633	0,6697
DOLP(6)	167,7500	0,6675
DOLP(7)	173,6883	0,6672
DOLP(8)	175,5740	0,6572
FSD(2)	3916,3753	0,5946
FSD(3)	1021,3813	0,5811
FSD(4)	175,8030	0,6003
FSD(5)	167,0920	0,6048
FSD(6)	247,5333	0,5915
FSD(7)	411,9333	0,6140
FSD(8)	820,4367	0,6048
ROLP(2)	3549,7277	0,6525
ROLP(3)	941,2347	0,6531
ROLP(4)	175,7673	0,6654
ROLP(5)	149,8900	0,6726
ROLP(6)	187,0333	0,6636
ROLP(7)	198,1153	0,6565
ROLP(8)	202,4503	0,6589
Gradiente(2)	4036,1433	0,8641
Gradiente(3)	1142,2820	0,8768
Gradiente(4)	218,5310	0,8681

Método	MSE	Tiempo
Gradiente(5)	207,0500	0,8896
Gradiente(6)	313,2000	0,8992
Gradiente(7)	514,8700	0,9056
Gradiente(8)	999,5100	0,9029
Morfológica(2)	3356,5333	1,9473
Morfológica(3)	1681,0537	2,0079
Morfológica(4)	1046,6467	2,0063
Morfológica(5)	1496,7167	2,0350
Morfológica(6)	2824,8033	2,0473
Morfológica(7)	3892,8600	2,0898
Morfológica(8)	5313,5467	2,1217
DWT(2)	9066,5867	0,6861
DWT(3)	5668,8367	0,6936
DWT(4)	2437,1573	0,7021
DWT(5)	484,2130	0,6876
DWT(6)	215,6727	0,6966
DWT(7)	200,4707	0,6884
DWT(8)	201,2860	0,6887
SIDWT(2)	7269,4900	0,8719
SIDWT(3)	4787,0767	1,1691
SIDWT(4)	2974,5590	1,4983
SIDWT(5)	1702,0927	1,8499
SIDWT(6)	901,1507	2,2591
SIDWT(7)	469,0210	2,7582
SIDWT(8)	260,9673	3,2556

Tabla 6: Resultados de la fusión YIQ con combinación basada en el gradiente.

Método	MSE	Tiempo
Media	1179,8400	0,5738
M. ponderada	927,0397	0,5848
DOLP(2)	196,0743	0,7758
DOLP(3)	95,1690	0,7737
DOLP(4)	78,1017	0,7879
DOLP(5)	73,8933	0,7835
DOLP(6)	72,6360	0,7834
DOLP(7)	72,4183	0,7786
DOLP(8)	72,3217	0,7663
FSD(2)	202,0043	0,7056
FSD(3)	100,9600	0,7024
FSD(4)	95,5810	0,7168
FSD(5)	114,8037	0,7165
FSD(6)	169,2617	0,7133
FSD(7)	323,2033	0,7102
FSD(8)	725,3867	0,7107
ROLP(2)	196,2227	0,7783
ROLP(3)	96,4803	0,7822
ROLP(4)	84,9797	0,7870
ROLP(5)	78,4180	0,7845
ROLP(6)	76,7087	0,7894
ROLP(7)	76,6603	0,7886
ROLP(8)	76,4563	0,7967
Gradiente(2)	219,8330	1,1181
Gradiente(3)	124,8767	1,1431
Gradiente(4)	128,3977	1,1471

Método	MSE	Tiempo
Gradiente(5)	150,7200	1,1524
Gradiente(6)	213,8067	1,1565
Gradiente(7)	373,7767	1,1466
Gradiente(8)	805,8133	1,1523
Morfológica(2)	3726,9767	2,0078
Morfológica(3)	3611,6733	2,0714
Morfológica(4)	7144,2133	2,0938
Morfológica(5)	12051,4667	2,1114
Morfológica(6)	30169,1800	2,1262
Morfológica(7)	19455,4600	2,1552
Morfológica(8)	23273,5533	2,1655
DWT(2)	474,2533	0,7807
DWT(3)	176,7060	0,7858
DWT(4)	98,5857	0,7835
DWT(5)	90,4587	0,7963
DWT(6)	77,7289	0,7942
DWT(7)	76,6091	0,7782
DWT(8)	77,6310	0,7830
SIDWT(2)	112,2510	1,1007
SIDWT(3)	112,8673	1,4442
SIDWT(4)	110,3053	1,8321
SIDWT(5)	100,5493	2,2527
SIDWT(6)	90,0493	2,7013
SIDWT(7)	83,6067	3,2499
SIDWT(8)	81,0273	3,8030

Tabla 7: Resultados de la fusión RGB con combinación basada en el gradiente.

Método	MSE	Tiempo
Media	2029,7160	0,2059
M. ponderada	1356,8343	0,2791
DOLP(2)	176,3033	0,8333
DOLP(3)	70,5313	0,8414
DOLP(4)	64,3159	0,8379
DOLP(5)	64,8093	0,8432
DOLP(6)	66,4698	0,8522
DOLP(7)	67,6345	0,8500
DOLP(8)	67,8445	0,8758
FSD(2)	178,0413	0,6477
FSD(3)	75,2807	0,6662
FSD(4)	78,6350	0,6530
FSD(5)	103,3173	0,6674
FSD(6)	168,5250	0,6680
FSD(7)	355,5633	0,6532
FSD(8)	820,7733	0,6528
ROLP(2)	188,3990	0,8830
ROLP(3)	71,6563	0,9292
ROLP(4)	68,5460	0,9336
ROLP(5)	66,9862	0,9170
ROLP(6)	66,8444	0,9074
ROLP(7)	67,3585	0,9062
ROLP(8)	67,9204	0,9047
Gradiente(2)	192,7127	1,9153
Gradiente(3)	92,4890	1,9830
Gradiente(4)	102,5653	2,0201

Método	MSE	Tiempo
Gradiente(5)	131,5093	1,9979
Gradiente(6)	196,6267	1,9743
Gradiente(7)	363,7767	2,0402
Gradiente(8)	815,7533	1,9690
Morfológica(2)	3971,1667	4,6229
Morfológica(3)	3257,4367	4,7566
Morfológica(4)	10881,5500	4,8443
Morfológica(5)	22577,4367	4,9128
Morfológica(6)	29521,6400	4,9653
Morfológica(7)	32606,9167	5,0693
Morfológica(8)	70939,8900	5,2509
DWT(2)	601,4943	0,8908
DWT(3)	143,5963	0,9048
DWT(4)	76,7993	0,9121
DWT(5)	82,6269	0,8987
DWT(6)	81,7487	0,9044
DWT(7)	85,8928	0,9122
DWT(8)	85,9933	0,9002
SIDWT(2)	139,7257	1,8905
SIDWT(3)	134,3017	2,9063
SIDWT(4)	124,5773	4,1374
SIDWT(5)	99,4217	5,4240
SIDWT(6)	89,3427	6,9089
SIDWT(7)	90,4693	8,4716
SIDWT(8)	95,9403	10,1554

Tabla 8: Resultados de la fusión de imágenes con dimensiones inadecuadas sin re-escalado.

Método	MSE	Tiempo
DOLP(2)	31,0890	0,0597
DOLP(3)	25,5050	0,0857
DOLP(4)	22,1515	0,1121
DOLP(5)	19,8842	0,1381
DOLP(6)	18,2351	0,1650
DOLP(7)	16,9825	0,1935
DOLP(8)	15,9985	0,2196
ROLP(2)	33,6597	0,0687
ROLP(3)	27,4010	0,0993
ROLP(4)	23,5357	0,1296
ROLP(5)	20,8924	0,1601
ROLP(6)	18,9503	0,1904
ROLP(7)	17,4854	0,2226
ROLP(8)	16,3215	0,2533
Gradiente(2)	31,8440	0,1974
Gradiente(3)	45,6430	0,2973
Gradiente(4)	60,3237	0,3959
Gradiente(5)	74,5050	0,4971
Gradiente(6)	88,0497	0,5828
Gradiente(7)	100,9623	0,6765
Gradiente(8)	113,1800	0,7691

Método	MSE	Tiempo
Morfológica(2)	68,1820	0,5928
Morfológica(3)	69,0040	0,8851
Morfológica(4)	72,0787	1,1779
Morfológica(5)	76,1510	1,4666
Morfológica(6)	80,4613	1,7542
Morfológica(7)	84,6500	2,0446
Morfológica(8)	88,5007	2,3356
DWT(2)	44,0027	0,1304
DWT(3)	37,0520	0,1939
DWT(4)	32,6063	0,2548
DWT(5)	29,4817	0,3185
DWT(6)	27,1377	0,3810
DWT(7)	25,2980	0,4461
DWT(8)	23,8243	0,5097
SIDWT(2)	28,7980	0,1449
SIDWT(3)	19,0215	0,2296
SIDWT(4)	16,2475	0,3336
SIDWT(5)	16,5087	0,4517
SIDWT(6)	18,1833	0,5830
SIDWT(7)	20,3913	0,7328
SIDWT(8)	22,6280	0,9024

Tabla 9: Resultados de la fusión de imágenes con dimensiones inadecuadas añadiendo bandas negras.

Método	MSE	Tiempo
DOLP(2)	14,3499	0,0670
DOLP(3)	9,9613	0,0703
DOLP(4)	9,9759	0,0749
DOLP(5)	10,3677	0,0745
DOLP(6)	10,5714	0,0751
DOLP(7)	10,9734	0,0772
DOLP(8)	11,2560	0,0755
FSD(2)	15,1945	0,0462
FSD(3)	12,0702	0,0489
FSD(4)	13,5841	0,0518
FSD(5)	15,5236	0,0523
FSD(6)	19,7606	0,0521
FSD(7)	30,2958	0,0540
FSD(8)	44,0171	0,0526
ROLP(2)	14,3973	0,0704
ROLP(3)	9,5682	0,0761
ROLP(4)	10,0440	0,0809
ROLP(5)	10,7625	0,0820
ROLP(6)	11,0586	0,0819
ROLP(7)	11,5471	0,0821
ROLP(8)	12,1304	0,0816
Gradiente(2)	19,9470	0,1344
Gradiente(3)	18,1265	0,1414
Gradiente(4)	21,3895	0,1524
Gradiente(5)	27,4100	0,1501

Método	MSE	Tiempo
Gradiente(6)	37,4747	0,1509
Gradiente(7)	55,3637	0,1507
Gradiente(8)	86,1457	0,1532
Morfológica(2)	87,9590	0,3866
Morfológica(3)	111,8880	0,4189
Morfológica(4)	146,7133	0,5648
Morfológica(5)	185,1967	0,5802
Morfológica(6)	277,6100	0,6002
Morfológica(7)	288,1100	0,6158
Morfológica(8)	464,1033	0,6310
DWT(2)	42,4577	0,0625
DWT(3)	28,1633	0,0679
DWT(4)	22,4453	0,0765
DWT(5)	21,0973	0,0765
DWT(6)	20,9567	0,0769
DWT(7)	21,0333	0,0776
DWT(8)	21,1690	0,0772
SIDWT(2)	28,7980	0,1464
SIDWT(3)	19,0215	0,2382
SIDWT(4)	16,2475	0,3526
SIDWT(5)	16,5087	0,4772
SIDWT(6)	18,1833	0,6168
SIDWT(7)	20,3913	0,7763
SIDWT(8)	22,6280	0,9585

Tabla 10: Resultados de la fusión de imágenes con dimensiones inadecuadas tras re-escalado por vecino próximo.

Método	MSE	Tiempo
DOLP(2)	14,8480	0,0722
DOLP(3)	10,5666	0,0760
DOLP(4)	10,5675	0,0796
DOLP(5)	10,9669	0,0797
DOLP(6)	11,1677	0,0799
DOLP(7)	11,5518	0,0803
DOLP(8)	11,8251	0,0803
FSD(2)	15,3873	0,0524
FSD(3)	12,5229	0,0546
FSD(4)	14,8630	0,0570
FSD(5)	19,1799	0,0578
FSD(6)	26,5202	0,0577
FSD(7)	39,3307	0,0579
FSD(8)	54,1767	0,0583
ROLP(2)	15,0683	0,0764
ROLP(3)	10,6288	0,0802
ROLP(4)	11,3426	0,0849
ROLP(5)	12,2283	0,0858
ROLP(6)	12,5427	0,0869
ROLP(7)	13,1506	0,0859
ROLP(8)	13,7870	0,0861
Gradiente(2)	20,1839	0,1396
Gradiente(3)	18,4607	0,1450
Gradiente(4)	21,8330	0,1554
Gradiente(5)	27,8830	0,1558

Método	MSE	Tiempo
Gradiente(6)	37,8657	0,1572
Gradiente(7)	55,2707	0,1573
Gradiente(8)	84,9803	0,1572
Morfológica(2)	86,9043	0,3896
Morfológica(3)	110,6103	0,4247
Morfológica(4)	141,6267	0,5687
Morfológica(5)	183,9500	0,5864
Morfológica(6)	258,3333	0,6025
Morfológica(7)	277,1200	0,6194
Morfológica(8)	396,9700	0,6359
DWT(2)	42,4153	0,0695
DWT(3)	28,4737	0,0733
DWT(4)	22,4933	0,0820
DWT(5)	21,3723	0,0817
DWT(6)	21,0893	0,0824
DWT(7)	21,2193	0,0832
DWT(8)	21,3630	0,0828
SIDWT(2)	28,9507	0,1516
SIDWT(3)	19,2286	0,2432
SIDWT(4)	16,3737	0,3572
SIDWT(5)	16,4617	0,4845
SIDWT(6)	17,9957	0,6203
SIDWT(7)	20,0563	0,7861
SIDWT(8)	22,2213	0,9602

Tabla 11: Resultados de la fusión de imágenes con dimensiones inadecuadas tras re-escalado bicúbico.

Método	MSE	Tiempo
DOLP(2)	47,5763	0,0987
DOLP(3)	39,9937	0,1011
DOLP(4)	40,4280	0,1053
DOLP(5)	40,8197	0,1056
DOLP(6)	41,0230	0,1061
DOLP(7)	41,3933	0,1061
DOLP(8)	41,6670	0,1061
FSD(2)	47,9140	0,0773
FSD(3)	41,8973	0,0804
FSD(4)	44,6337	0,0848
FSD(5)	48,9503	0,0837
FSD(6)	56,3207	0,0841
FSD(7)	69,1820	0,0839
FSD(8)	84,1370	0,0853
ROLP(2)	48,0893	0,1021
ROLP(3)	39,7290	0,1062
ROLP(4)	40,6093	0,1116
ROLP(5)	41,4257	0,1117
ROLP(6)	41,7573	0,1116
ROLP(7)	42,2237	0,1122
ROLP(8)	42,7350	0,1123
Gradiente(2)	53,7230	0,1660
Gradiente(3)	48,6320	0,1723
Gradiente(4)	52,2557	0,1823
Gradiente(5)	58,1900	0,1813

Método	MSE	Tiempo
Gradiente(6)	68,0973	0,1835
Gradiente(7)	85,4487	0,1845
Gradiente(8)	115,1573	0,1827
Morfológica(2)	101,6987	0,4154
Morfológica(3)	119,3407	0,4492
Morfológica(4)	147,4233	0,5931
Morfológica(5)	179,4067	0,6100
Morfológica(6)	231,2667	0,6281
Morfológica(7)	241,6167	0,6436
Morfológica(8)	325,2300	0,6592
DWT(2)	74,2233	0,0950
DWT(3)	57,0050	0,1004
DWT(4)	51,7763	0,1082
DWT(5)	50,6440	0,1082
DWT(6)	50,3750	0,1091
DWT(7)	50,5100	0,1085
DWT(8)	50,6760	0,1086
SIDWT(2)	63,6980	0,1765
SIDWT(3)	50,8480	0,2713
SIDWT(4)	49,0540	0,3865
SIDWT(5)	49,5213	0,5078
SIDWT(6)	51,5250	0,6510
SIDWT(7)	53,8840	0,8061
SIDWT(8)	56,4190	0,9880

Bibliografía

- Adelson, E. H., Anderson, C. H., Bergen, J. R., Burt, P. J., & Ogden, J. (1984). Pyramid methods in image processing. *RCA Eng* , 9 (6), 33–41.
- Amolins, K., Zhang, Y., & Dare, P. (2007). Wavelet based image fusion techniques – An introduction, review and comparison. *Journal of Photogrammetry & Remote Sensing* , 62, 249-263.
- Anderson, C. H. (1987). *Patent No. 4.718.104*. United States.
- Anderson, C. H. (1984). An alternative to the Burt pyramid algorithm. *RCA correspondance* .
- Beylkin, G. (1992). On the representation of operators in bases of compactly supported wavelets. *SIAM J. Numer. Anal.* , 29 (6), 1716-1740.
- Burrus, C. S., Gopinath, R. A., & Guo, H. (1998). *Introduction to Wavelets and Wavelet Transforms: a Primer*. Upper Saddle River: Prentice-Hall.
- Burt, P. (1992). A gradient pyramid basis for pattern selective image fusion. *The Society for Information Displays (SID) International Symposium Digest of Technical Papers* , 23, 467-470.
- Burt, P. J. (1984). The pyramid as a structure for efficient computation. *Multiresolution Image Processing and Analysis* , 6-35.
- Burt, P. J., & Adelson, E. H. (1983). The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications* , 31 (4), 532-540.
- Burt, P. J., & Kolczynski, R. J. (1993). Enhanced Image Capture Through Fusion.
- Burt, P. J., van der Wal, G. S., Kolczynski, R. J., & Hingorani, R. (1993). *Patent No. 5.488.674*. United States.
- Equipo de desarrollo de Gimp. (n.d.). *Gimp for Windows*. Retrieved 2010, from <http://gimp-win.sourceforge.net/>
- Gonzalez, R. C., & Woods, R. E. (1993). *Digital Image Processing*. Boston: Addison-Wesley.
- Haeberli, P. (1994). *A Multifocus Method for Controlling Depth of Field*. Retrieved from <http://www.graficaobscura.com/depth/index.html>
- Haralick, R. M., Lin, C., Lee, J. S., & Zhuang, X. (1987). Multiresolution morphology. *Proceedings of IEEE First International Conference on Computer Vision* , 516-520.
- Heijmans, H. J., & Goutsias, J. (2005). Morphological Decomposition Systems with Perfect Reconstruction: From Pyramids to Wavelets. *Space, Structure and Randomness* , 183, 279-314.
- Intel. (n.d.). *Intel® Core™ i5-430M Processor (3M Cache, 2.26 GHz)*. Retrieved 2010, from <http://ark.intel.com/Product.aspx?id=43537>
- Li, H., Manjunath, B. S., & Mitra, S. K. (1994). Multisensor Image Fusion Using the Wavelet Transform. *Graphical Models and Image Processing* , 57, 235-245.

Maik, V., Cho, D., Shin, J., & Paik, J. (2007). Regularized Restoration Using Image Fusion for Digital Auto-Focusing. *IEEE Transactions on Circuits and Systems for Video Technology* , 17 (10).

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* , 11, 674–693.

Mertens, T., Kautz, J., & Van Reeth, F. (2007). Exposure Fusion.

Microsoft Corporation. (n.d.). *Microsoft Windows 7*. Retrieved 2010, from <http://www.microsoft.com/windows/windows-7/>

Ogden, J. M., Adelson, E. H., Bergen, J. R., & Burt, P. J. (1985). Pyramid-based computer graphics. *RCA Engineer* , 30 (5), 4-15.

Pajares, G., & de la Cruz, J. M. (2004). A wavelet-based image fusion tutorial. *Pattern Recognition* , 7, 1855-1872.

Pajares, G., & de la Cruz, J. M. (2001). *Visión por Computador: Imágenes Digitales y Aplicaciones*. Madrid: RA-MA.

Ramac, L. C., Uner, M. K., & Varshney, P. K. (1998). Morphological filters and wavelet based image fusion for concealed weapon detection. *Proceedings of SPIE* , 3376.

Rockinger, O. (1999, septiembre 30). *image fusion*. Retrieved from <http://www.metapix.de/examples.htm>

Rockinger, O. (1997). Image Sequence Fusion Using a Shift-Invariant Wavelet Transform. *IEEE Transactions on Image Processing* , 3, 288–291.

Rosendfeld, A., & Thurston, M. (1971). Edge and curve detection for visual scene analysis. *IEEE Transactions on Computers* , 20, 562–569.

Sadjadi, F. (2005). Comparative Image Fusion Analysis. *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 3, pp. 20-26.

Santos, M., Pajares, G., Portela, M., & de la Cruz, J. M. (2003). A New Wavelets Image Fusion Strategy. *Lecture Notes in Computer Science* , 2652, 919-926.

Sims, O., & Irvine, J. (2006). An FPGA Implementation of Pattern-Selective Pyramidal Image Fusion. *Field Programmable Logic and Applications* .

The MathWorks. (1984-2009). Converting Color Data Between Color Spaces. In *Matlab User Guide*.

The MathWorks. (n.d.). *MATLAB - The Language Of Technical Computing*. Retrieved 2010, from <http://www.mathworks.com/products/matlab/>

TheWebfoto. (2008). *Profundidad de campo - Curso de fotografía digital*. Retrieved from <http://www.thewebfoto.com/2-hacer-fotos/211-profundidad-de-campo/>

Toet, A. (1990). Hierarchical image fusion. *Machine Vision and Applications* , 3, 1-11.

Toet, A. (1989). Image fusion by a ratio of low-pass pyramid. *Pattern Recognition Letters* , 9, 245-253.

Toet, A., van Ruyven, J. J., & Valetton, J. M. (1989). Merging thermal and visual images by a contrast pyramid. *Optical Engineering* , 28 (7), 789-792.

Tsai, D. M., & Hsiao, B. (2001). Automatic surface inspection using wavelet reconstruction. *Pattern Recognition* , 34, 1285–1305.

Yocky, D. A. (1995). Image merging and data fusion by means of the discrete two-dimensional wavelet transform. *Journal of the Optical Society of America A* , 12 (9), 1834–1841.

Zhang, Z., & Blum, R. S. (1999). A categorization of Multiscale-Decomposition-Based Image Fusion Schemes with a Performance Study for a Digital Camera Application. *Proc. IEEE* , 87 (8), 1315–1326.

Zitová, B., & Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing* , 21, 977-1000.

Autorización

El abajo firmante, matriculado en el Máster en Investigación en Informática de la Facultad de Informática, autoriza a la Universidad Complutense de Madrid (UCM) a difundir y utilizar con fines académicos, no comerciales y mencionando expresamente a su autor el presente Trabajo Fin de Máster: “Análisis de técnicas de fusión de imágenes para su uso en fotografía digital”, realizado durante el curso académico 2009-2010 bajo la dirección de Gonzalo Pajares Martinsanz en el Departamento de Arquitectura de Computadores y Automática, y a la Biblioteca de la UCM a depositarlo en el Archivo Institucional E-Prints Complutense con el objeto de incrementar la difusión, uso e impacto del trabajo en Internet y garantizar su preservación y acceso a largo plazo.

En Madrid, a 14 de junio de 2010

Daniel Alonso Fernández